INSTITUTE of HYDROLOGY

AN EVALUATION OF FLOW FORECASTING PROCEDURES FOR THE CITARUM RIVER BASIN, INDONESIA

by

P.E. O'CONNELL

(Institute of Hydrology)

INSTITUTE OF HYDROLOGY
WALLINGFORD
OXON.

SEPTEMBER 1981

SUMMARY

The purpose of the consultant's mission was to carry out an independent evaluation of the forecasting procedures implemented under UNDP/WMO Project INS/78/038 for the Citarum River Basin. The forecasting system developed under the project, which utilizes the COSSARR model for forecasting flows and reservoir levels up to two days ahead, underwent its first 'trial run' during the 1980/81 high water season, and is scheduled to have its first operational run in 1981/82.

The consultant's approach to the evaluation of forecasting procedures was to calibrate some simple models for the flows at two gauging stations within the basin, to use these models for forecasting over the 'trial run' period and to compare the results with those obtained for the COSSARR model. Prior to this work, a review of the basic rainfall and flow data used to calibrate the COSSARR model was carried out. Since it was found that a significant proportion of the rainfall data had been infilled, rainfall stations with complete records of observed data were identified and further COSSARR calibration studies were initiated together with the calibration of some simple models. The results showed that the agreement between observed and simulated flows obtained with the simple models was similar to that obtained with the more complex COSSARR model.

Over the 1981/82 'trial-run' forecast period, forecasts of rainfall one and two days ahead was made within the Project using qualitative meteorological information coupled with a quantitative statistical procedure; a comparison with other procedures showed that these forecasts are as good as can be obtained with the available information. A comparison of one day and two day ahead forecasts made by the COSSARR and simple models over the 'trial-run' period showed that the latter model gave slightly better forecasts.

The main conclusion drawn from the consultant's work is that, while the COSSARR model has performed adequately during its trial run, a similar forecasting accuracy (and hence level of benefits) can be obtained from a simpler model costing much less to implement and run operationally. If other forecasting projects are to be set up throughout Indonesia, then the use of simple forecasting models should be considered, since it is unlikely that the full capability of the COSSARR model to simulate river regulation by complex systems of reservoirs would be required.

ACKNOWLEDGEMENTS

The work described in this report could not have been carried out without the excellent cooperation and support provided by the personnel in Project INS/78/038 and by DPMA counterparts. In particular, the efforts of Mr Ohn Maung in preparing all the basic information for the evaluation of forecasting procedures is gratefully acknowledged, while Ir. Wanny Adidarma provided invaluable help on all computing aspects of the work. The administrative support provided by Mr. Surin Sangsnit was much appreciated. Finally, the consultant would like to thank the Director of DPMA and the staff of the Hydrology, Hydrometry and Computing Sections of DPMA for their help and assistance.

APPENDICES

1. BACKGROUND AND TERMS OF REFERENCE

BACKGROUND AND TERMS OF REFERENCE

River Forecasting Project INS/78/038, under which a forecasting system has been developed and implemented for the Citarum River Basin, is now entering its third year; the system, which employs the COSSARR model for forecasting flows and reservoir levels, underwent its first 'trial run' during the period December 1980 - April 1981, and is scheduled to provide 'operational run' forecasts for the 1981-82 high water season.

The appointment as Consultant to Project INS/78/038 was made under WMO Special Service Agreement No. 29.743/A/PEX (dated 11 June 1981) for the period 6 July - 22 Aug 1981, with exclusion of the period 1-11 Aug. The terms of reference for the appointment were specified in the Special Service Agreement as follows:

> 'to prepare an independent evaluation of flood forecasting procedures and of the usefulness of the various forecasting models implemented under the project'.

The consultant arrived in Jakarta on 7 July, and visited the Meteorological and Geophysical Institute and UNDP offices in Jakarta on 8 July. The periods 9 - 31 July and 12 - 18 Aug were spent at the Project Office in DPMA, Bandung where the Programme of Work described in Section 2 of this report was carried out; the consultant arrived in Geneva on 20 August and visited WMO headquarters on 21 August, departing for London on 22 August.

2. PROGRAMME OF WORK

PROGRAMME OF WORK

During the consultant's assignment, work was carried out under the following headings:

(a) Review of basic data: the rainfall and flow data used to calibrate the various models implemented under the project were reviewed to assess if these were of satisfactory quality;

(b) Calibration of simple flow forecasting models : the COSSARR model has been calibrated for the Citarum River Basin using daily rainfall and flow data for the period 1974-77; a number of simpler models were calibrated during the consultant's mission to allow comparisons with the results obtained from the COSSARR model;

(c) Evaluation of results for 'trial run' forecast period, Dec.1980-Apr. 1981:

the COSSARR model was used to provide one day and two day ahead forecasts of flow in the Citarum River and of Jatiluhur reservoir level over the above period; one and two day ahead forecasts of rainfall were also required for this purpose. A number of error statistics have been calculated for these forecasts and compared with those obtained from some of the simpler models calibrated under (b) above;

(d) Lectures: three lectures were delivered on the following topics:

1. Raingauge network rationalization
2. Rainfall-runoff modelling
3. Real-time flow forecasting

Summaries of these lectures are given in Appendix A;

(e) Preparation of report: this report describes the programme of work carried out by the consultant

3. THE FORECASTING SYSTEM FOR THE CITARUM RIVER BASIN

3. THE FORECASTING SYSTEM FOR THE CITARUM RIVER BASIN

River Forecasting Project INS/78/038 commenced in August 1979 with the objective of establishing a forecasting system for the Citarum River Basin which would act as a pilot project for possible future river forecasting projects throughout Indonesia. In developing the forecasting system, work has been carried out under the following major headings:

(a) installation of reporting network

(b) calibration of COSSARR model for forecasting river flows and reservoir levels;

(c) statistical studies of rainfall characteristics and patterns;

(d) operational testing of forecasting system during 'trial-run' period December 1980 - April 1981.

In the first year of the project, a network of single side band (SSB) radio transmitters was installed; the observers at these stations, of which there are 11 distributed throughout the basin, report directly to the Project Office at DPMA. (Figure 3.1). In addition, there are a number of other rainfall reporting stations which transmit daily rainfall amounts through other channels of communication (Figure 3.1).

The SSARR river basin model, and its derivative, the COSSARR model (a version of SSARR developed for relatively small basins and small computers) are fully documented in the Project reports (e.g. Rockwood, 1980; Sangsnit, 1980) and will not be described here. In preparation for operational usage, the COSSARR model was calibrated for three areas (Rockwood, 1980).

(a) the Citarum River at Nanjung (area 1718 km^2)

(b) the Citarum River at Palumbon (area 4061 km^2)

(c) Palumbon Local (area 4061-1718 = 2343 km^2)

For (c) the Palumbon Local inflows were obtained for the years 1974-77 by routing the observed flows at Nanjung through the channel storage from Nanjung to Palumbon and subtracting the routed flows from the observed flows at Palumbon. Thus, simulated flow at Palumbon is obtained by simulating the flow contribution from the Palumbon Local area and adding this to the simulated flow at Nanjung routed to Palumbon. The data

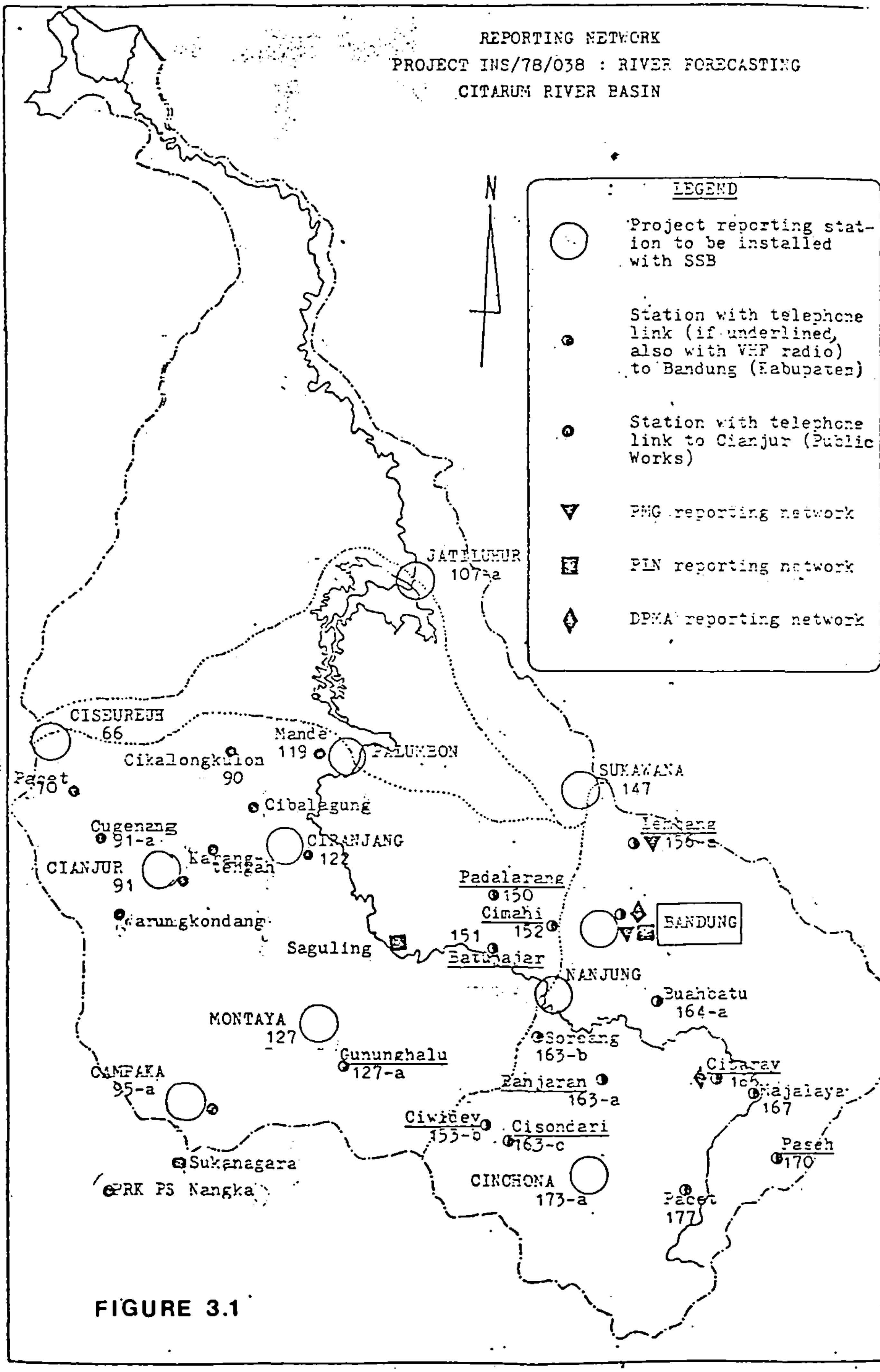

REPORTING NETWORK
PROJECT INS/78/038 : RIVER FORECASTING
CITARUM RIVER BASIN
N
LEGEND
Project reporting station to be installed with SSB
Station with telephone link (if underlined, also with VHF radio) to Bandung (Kabupaten)
Station with telephone link to Cianjur (Public Works)
PMG reporting network
PLN reporting network
DPMA reporting network
JATILUHUR 107-a
CISEUREUH 66
Mande 119
Cikalongkulon 90
Pacet 70
Cibalagung
SUKAWANA 147
Cugenang 91-a
CIRANJANG 122
CIANJUR 91
Padalarang 150
Cimahi 152
151
BANDUNG
Saguling
Batujajar
NANJUNG
Buahbatu 164-a
MONTAYA 127
Gununghalu 127-a
163-b
Banjaran 163-a
Majalaya 167
Ciwidey
Cisondari 163-c
Paseh 170
Sukanagara
PRK PS Nangka
CINCHONA 173-a
Pacet 177

FIGURE 3.1

used in these calibration studies are described in Section 4, and the results obtained are discussed in Section 5.

The statistical studies of rainfall characteristics and patterns in the Citarum River Basin have been carried out to assist in making one day and two day ahead forecasts of rainfall during the 'trial run' forecast period December 1980 - April 1981, as no quantitative meteorological information is available to the Project in real-time other than the general forecast from the Indonesian Meteorological and Geophysical Institute (BMG) that rainfall over the Citarum River Basin in the coming 24 hours will be either

(a) Isolated,

(b) Scattered,

or (c) Widespread

Using historical rainfall data for a number of stations in the Citarum River Basin, the frequencies (in four classes) of basin rainfall have been computed by Sangsnit and Maung (1981). Thus, the BMG forecast identifies the type of rainfall to be expected; then if, for example, the rainfall in the previous part of the month has been above average, one of the upper classes for the appropriate category is used as the forecast.

In making one day and two day ahead forecasts of discharge over the period December 1980 - April 1981, a three hourly time step was used by the COSARR model. At 7.00 each morning, the SSB network relays to the Project Office rainfall amounts at SSB stations during the previous 24 hours, and stage levels for Nanjung, Palumbon and Jatiluhur reservoir. Using these daily rainfall values augmented with an isohyetal map of monthly rainfall, isohyets of basin rainfall are drawn by hand and basin averages computed for the Nanjung, Palumbon Local and Jatiluhur Local areas. If data from some of the other reporting stations become available in time, these are included. The basin averages, thus computed, for the current day and two previous days (the 'back-up' period) are then fed into the COSSARR model together with forecasted basin rainfalls for the coming two days (the forecast period) and processed in one operation to give one day and two day ahead forecasted flows and reservoir levels at 07.00 hours. The observed basin rainfalls during the 'back-up' period are adjusted iteratively until the model is deemed to accurately represent the observed flows in this period prior to making forecasts. Forecasted daily basin rainfalls are broken down into three hourly totals in accordance with the fairly regular observed distribution of rainfall in time; in the 'back-up' period, this distribution may be altered if information to this effect has been received. Execution time on the IBM 1130 at DPMA takes about 20 minutes; forecasts are disseminated to the relevant authorities by 12 noon on each day.

The results obtained during the 'trial-run' forecast period Dec. 1980 - Apr. 1981 are discussed in Section 6.

4. REVIEW OF BASIC DATA

4. REVIEW OF BASIC DATA

4.1 Rainfall data

In a previous report, Sugawara (1980) had suggested that the rainfall data supplied for calibration of the Tank Model on the Citarum River Basin were not observed data; the daily rainfall data at some stations had apparently been infilled from those at others. Sugawara identified a number of stations for which he considered observed data were available and used these in calibration studies of the Tank Model. Through reference to the original manuscripts obtained from the Meteorological and Geophysical Institute (BMG), a list of rainfall stations was drawn up for which observed data were available for the period 1974-76 (Table 4.1); most of these stations lie within the Citarum River Basin. This list was then used to check the rainfall data used in the COSSARR and Tank Model calibration studies; for stations where the daily data had been infilled, the stations used to do the infilling were also identified.

The rainfall stations used in the COSSARR calibration runs are listed in Table 4.2 for Nanjung and Palumbon together with the weights used in computing average daily basin rainfall. These weights are taken as the ratios of annual average basin rainfall to annual average rainfall at the stations in question; two sets of stations are listed since it was found that the final COSSARR calibration runs had used additional rainfall stations. In Table 4.3, these rainfall data are classified into observed and infilled, and the relationships existing between them are also depicted. The infilling had been carried out using monthly scaling factors which were taken as the ratios of the long-term average monthly rainfalls at the stations in question. Since daily rainfall totals vary greatly over the Citarum River Basin due to the localized nature of rain storms, this infilling procedure cannot be considered satisfactory, particularly for daily rainfall-runoff modelling.

From Tables 4.2 and 4.3, it can be seen that the estimates of average basin rainfall for Nanjung and Palumbon Local basins used in the early and final COSSARR calibration runs were based on observed data from a relatively small number of stations. While it is appreciated that the infilled stations were chosen to coincide as far as possible with the stations in the reporting network; this in itself does not constitute a sufficient basis for selection, particularly when, in the case of the Nanjung basin, relatively few stations with observed data were used for infilling.

TABLE 4.1 Stations for which observed rainfall data available for period January, 1974-December, 1976

66	CISEUREUH	136	CICACING
70	PACET	145a	CIMANGSUD (PERK)
77	LAMPEGAN (PERK HARJISARI)	147	SUKAWANA
90a	VADA	150	PADALARANG
91a	CUGENANG	151c	BATUJAJAR
94	GUNUNG CEMPAKA	153b	CIWIDEY
95a	CAMPAKA	154b	MARGAHAYU
98	SUKANEGARA	156a	LEMBANG
122	CIRANJANG	160	PAKAR
123	PASIK GOMBONG	163	BANDUNG
125	BOJUNGPICUNG	163c	CISONDARI
125a	BOJUNGPICUNG (PERK)	164	JATINANGOR (PERK)
126	CIBARENGKOK	168	ARJASARI (PERK)
127	MONTAYA	170	CIDAKU (PASEH)
127a	GUNUNG HALU	174	CIBEUREUM
		185	PERK JALUN

Table 4.2 Lists of rainfall stations used to calculate basin rainfalls for COSSARR calibration runs

(a) Early COSSARR Calibration Runs

Nanjung	Weight	Palumbon Local	Weight
162a CIHEMPELAS(I)	0.95	66 CISEUREUH(0)	0.72
163c CISONDARI (0)	0.90	91 CIANJUR (I)	0.99
167 MAJALAYA (I)	1.17	94 GUNUNG CAMPAKA(0)	0.97
172 CINYIRUAN(I)	0.77	127 MONTAYA (0)	0.96
		151a SINDANGKERTA(I)	0.97
		162a CIHEMPELAS (I)	1.18

(b) Final COSSARR calibration runs

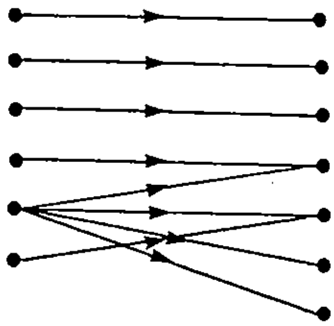
Nanjung	Weight	Palumbon Local	Weight
160 PAKAR(0)	0.99	66 CISEUREUH (0)	0.72
172a CIHEMPELAS(I)	0.95	91 CIANJUR (I)	0.99
163c CISONDARI (0)	0.90	94 GUNUNG CAMPAKA(0)	1.18
167 MAJALAYA (I)	1.17	151a SINDANGKERTA(I)	0.97
170 PASEH(0)	0.68	162a CIHEMPELAS(I)	1.18
172 CINYIRUAN(I)	0.77		
180 MALABAR (I)	0.87		

Note: 0 denotes observed while I denotes infilled

Table 4.3 Stations with observed and infilled rainfall data, and linkages between them

OBSERVED

90a VADA
127 MONTAYA
151c BATUJAJAR
160c PAKAR
163c CISONDARI
170 PASEH



INFILLED

91 CIANJUR
151a SINDANGKERTA
152 CIMAHI
162a CIHEMPELAS
167 MAJALAYA
172 CINYIRUAN
180 MALABAR

Table 4.4 Stations used by Sugawara (1980) in calculating average rainfall for Nanjung and Palumbon basins

Nanjung	Weight	Palumbon(total)	Weight
152 CIMAHI(I)	1.60	152 CIMAHI(I)	1.5
163c CISONDARI(O)	0.85	163c CISONDARI(O)	0.9
160 PAKAR(O)	1.35	160 PAKAR(O)	1.35
170 PASEH (O)	1.15	170 PASEH(O)	1.2
		66 CISEUREUH(O)	0.75
		91 CIANJUR (I)	1.00
		127 MONTAYA(O)	1.20
		90a VADA(O)	1.10
		150 PANDALARANG(O)	1.45

Note: O denotes observed while I denotes infilled.

The rainfall stations used by Sugawara (1980) for the Nanjung and Palumbon basins are listed in Table 4.4; of the total of 9 rainfall stations used to model flow at Palumbon the data for 7 were observed while those for 2 had been infilled.

Since it seemed desirable to identify a new set of rainfall stations for further model calibration studies from the full set of stations with observed data listed in Table 4.1, seven stations were identified within each of the Nanjung and Palumbon Local Basins; the locations were chosen to give good areal coverage and to be as near as possible to stations in the reporting network. The number of stations was limited to seven in each case since this is the maximum number that the COSSARR Program on the IBM 1130 computer at DPMA can handle; however, during subsequent model calibration studies, errors in simulating flow for the Palumbon local basins with other models were found to be attributable to the use of an insufficient number of stations in estimating average basin rainfall and so a further 3 stations were added to make a total of 10 stations for the latter basin. The weights used in computing average daily basin rainfall were calculated as described above.

Thus, to summarize, results for model calibration studies based on four different estimates of average basin rainfall will be presented in the report; these will be referred to as sets SA1 and SA2 corresponding to the initial and final COSSARR runs (results of model calibration studies for set SA1 will be presented since it was not established until the middle of the consultant's visit that set SA2 had been used in the final COSSARR runs), and sets NS1 and NS2 corresponding to the new selections described above; the new selections are summarized in Table 4.5. The locations of the rainfall stations used in selections SA1 and SA2 are shown in Figures 4.1 and 4.2 while Figure 4.3 shows the locations of the stations used in selections NS1 and NS2.

Table 4.5 New selections of rainfall stations for computing Nanjung and Palumbon Local basin rainfalls

(a) New selection NS1

Nanjung		Weight	Palumbon Local		Weight
156a	LEMBANG	1.09	66	CISEUREUH	0.75
163	BANDUNG	1.22	91a	CUGENANG	1.00
163c	CISONDARI	1.01	122	CIRANJANG	1.13
164	P.JATINANGOR	1.20	94	GUNUNG CAMPAKA	0.78
168	ARJASARI	0.90	127	MONTAYA	0.96
170	PASEH	0.97	147	SUKAWANA	0.96
185	PERK JALUN	0.87	151c	BATUJAJAR	1.36

(b) Extra stations added to selection NS1 to give selection NS2

Palumbon Local	Weight
77 PERK. HARJASARI	0.80
125 BOJUNGPICUNG	0.68
145a CIMANGSUD (PERK)	0.78

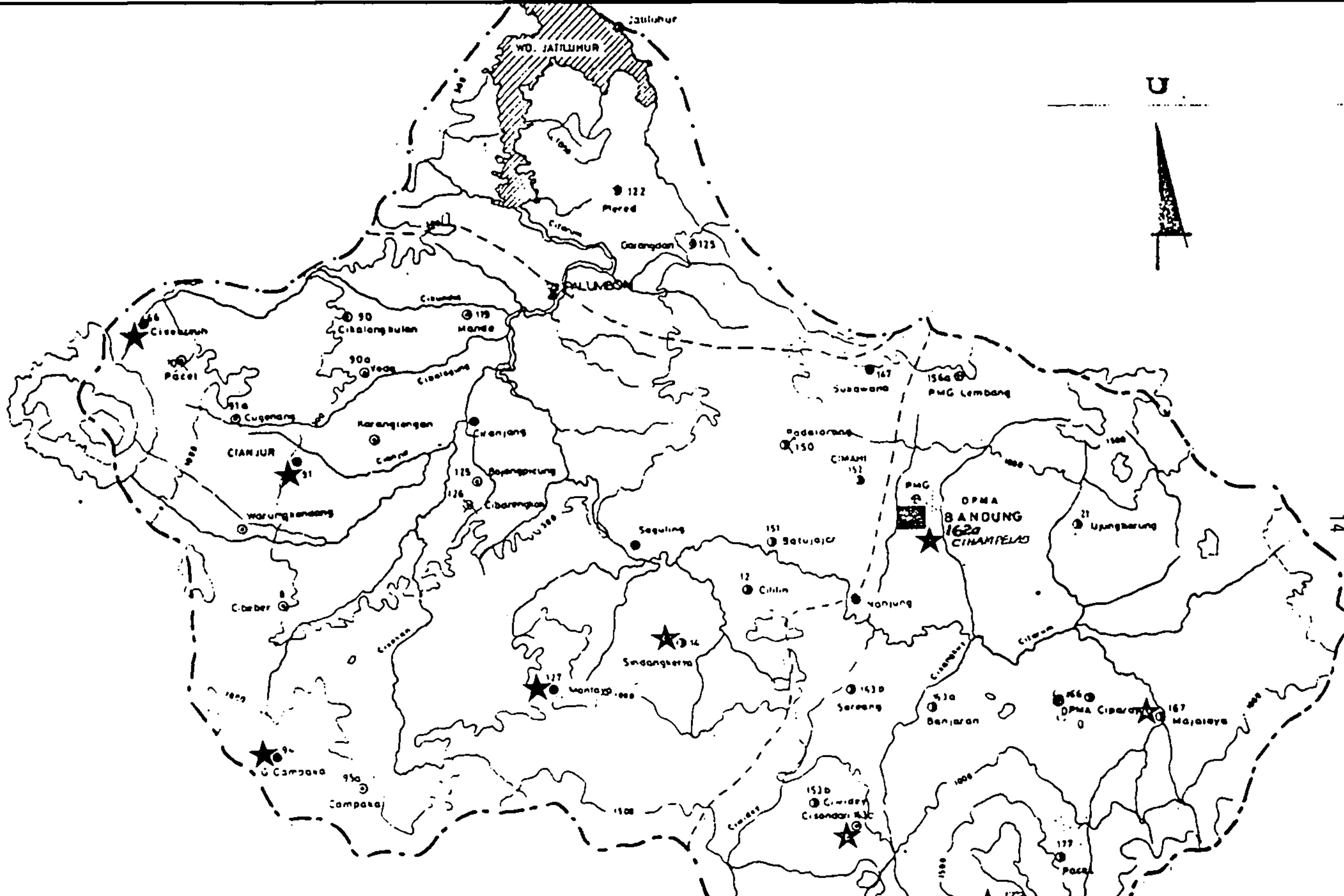
4.2 Flow data

The records of average daily discharge at Nanjung and Palumbon used in the COSSARR model calibration studies have been computed by visually assessing the average stage over each 24 hour period (midnight to midnight) and then converting these values to average discharge using the available rating curve. For days on which large fluctuations in stage occur, more accurate estimates of average daily discharge could be obtained by extracting a sufficient number of stage readings to define reasonably well the fluctuations in stage within the 24 hour period, converting these readings to discharge, and then averaging the resulting values.

As part of the Indonesian Floods Study currently being carried out jointly by DPMA and the Institute of Hydrology, UK, the rating curves for Nanjung and Palumbon have been assessed. In the case of Nanjung, a fair amount of scatter is observed at high flows; the maximum stage at which a gauging has been carried out is 4.3 m, while the maximum observed stages are 5.24 m (1931) and 5.03 m(1975). For Palumbon, there is little scatter up to the maximum gauging of 4.58 m ; the maximum observed stages are 8.52 m (1940) and 7.80 m (1978).

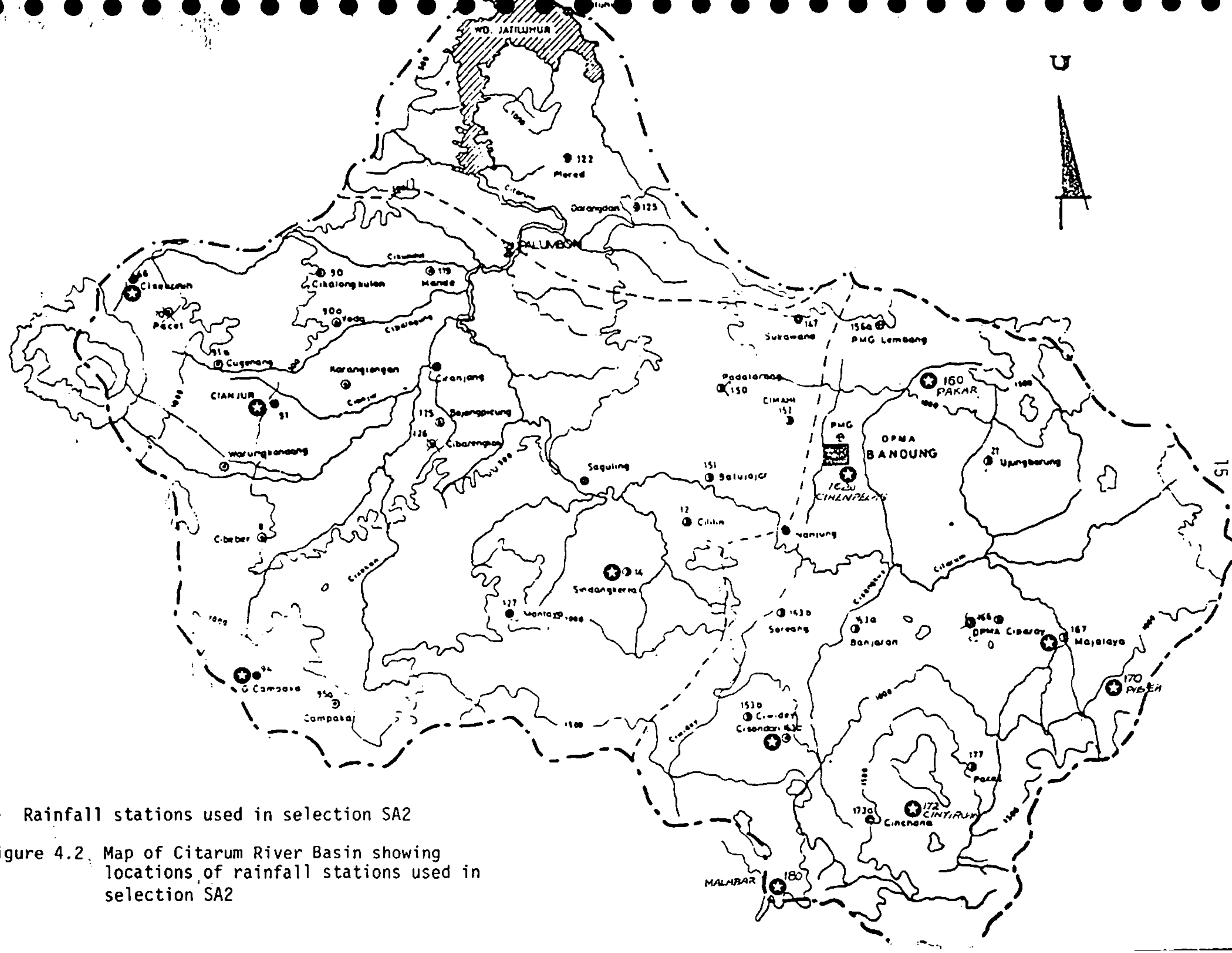
Within the Hydrometry Section at DPMA, rating curves are changed periodically on the basis that, if new gaugings depart from existing curves, the cross-sections may have changed. For example during the period 1974-76, the rating curve for Nanjung was found to have been changed in early 1976.

The discharge record at Nanjung for the period 1974-76 is complete; for Palumbon, the data for the period Nov.1 - Dec. 1, 1975 are missing.



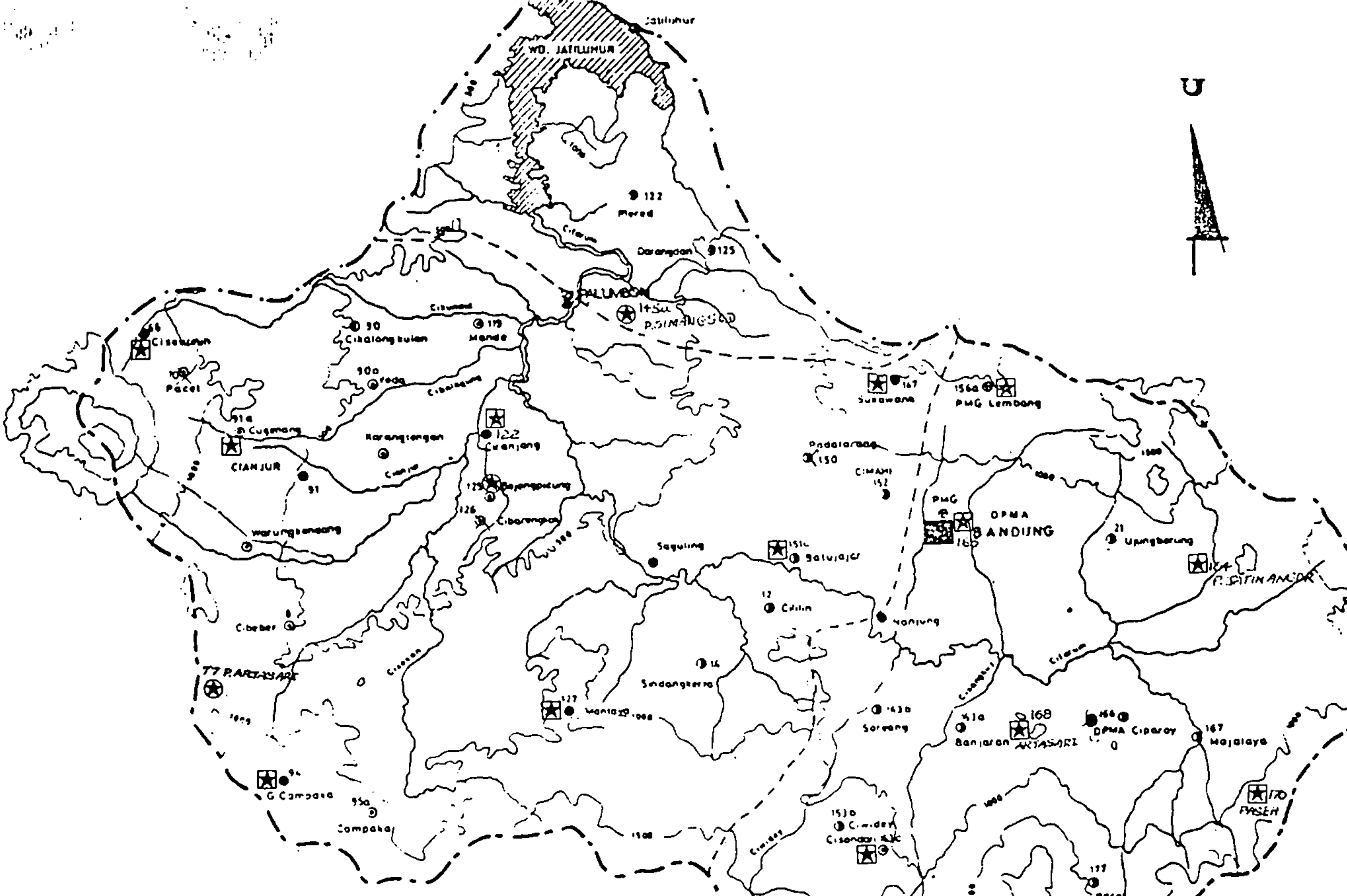
★ Rainfall stations used in selection SA1

Figure 4.1: Map of Citarum River Basin showing locations of rainfall stations used in selection SA1



Rainfall stations used in selection SA2

Figure 4.2. Map of Citarum River Basin showing locations of rainfall stations used in selection SA2



- Rainfall stations used in selection NS1
- Rainfall stations added to selection NS1 to give selection NS2

Figure 4.3 Map of Citarum River Basin showing locations of rainfall stations used in selections NS1 and NS2

5. CALIBRATION OF SIMPLE FLOW FORECASTING MODELS

CALIBRATION OF SIMPLE FLOW FORECASTING MODELS

5.1 Description of models

To facilitate an evaluation of the results obtained from the CO-SSARR model calibration studies carried out using data for the period 1974-77, a number of models with relatively simple structures were calibrated using historical flow data for Nanjung and Palumbon; as discussed in Section 4, results for 4 different estimates of average basin rainfall were obtained, all for the period January 1974-December 1976. The models which were fitted fall into 3 classes:

(i) linear and non-linear transfer function (TF) models;

(ii) constrained linear system (CLS) models;

(iii) simple conceptual models.

Brief descriptions of these models and the procedures used to calibrate them are given here; more detailed descriptions, and appropriate references are given in Appendix B.

For the basic linear TF model, it is assumed that observed discharge can be represented as

$$q_t = \mathit{q}_t + n_t \qquad (5.1)$$

where q_t is a deterministic component of flow and n_t is a stochastic component; the linear TF model is then used to represent q_t as

$$\mathit{q}_t = -\delta_1 \mathit{q}_{t-1} - \delta_2 \mathit{q}_{t-2} - \cdots - \delta_r \mathit{q}_{t-r} + \omega_0 p_{t-b} + \omega_1 p_{t-b-1} + \cdots \omega_{s-1} p_{t-b-s-1} \qquad (5.2)$$

where p_{t-b}, p_{t-b-1}, ... $p_{t-b-s-1}$ are rainfall inputs lagged by a pure time delay b, and δ_1 ... δ_r and ω_0 ...ω_{s-1} are r autoregressive and s moving average parameters, respectively; the model in shorthand notation may be written as TF(r,s,b). TF(r,s,b) models may be shown to be equivalent to the traditional impulse response or unit hydrograph representations of catchment response widely used in hydrology (Appendix B);

TF representations are however to be preferred on the grounds that they involve far fewer parameters.

In modelling the rainfall-runoff process, the assumption of linearity may prove restrictive; non linear TF(r,s,b) models can be obtained either by applying a transformation or a threshold to the rainfall input. If an antecedent precipitation index (API) is computed at time t as

$$API_t = K\,API_{t-1} + p_{t-1} \tag{5.4}$$

where K may be constant or may vary seasonally for daily data as

$$K_t = \bar{K} + \alpha \operatorname{Cos} \frac{2\pi(t-\phi)}{365}. \tag{5.5}$$

A transformed rainfall input can then be obtained as

$$p_t^* = API_t . p_t \tag{5.6}$$

whence p_t^* is used instead of p_t in equation (5.2). TF models with thresholds may be obtained by using the API to generate two separate rainfall inputs as follows:

$$\begin{aligned} API_t &> T \quad p_t^{(1)} = 0 \; ; \quad p_t^{(2)} = p_t \\ API_t &\leq T \quad p_t^{(1)} = p_t ; \quad p_t^{(2)} = 0. \end{aligned} \tag{5.7}$$

The TF model is then written as

$$q_t = - \delta_1 q_{t-1} - \dots - \delta_r q_{t-r} + \sum_{i=1}^{2} \omega_0^{(i)} p_{t-b}^{(i)} + \omega_1^{(i)} p_{t-b-1}^{(i)} \\ \dots \omega_{s-1}^{(i)} p_{t-b-s-1}^{(i)} \tag{5.8}$$

Procedures for the identification of values of r, s and b for a particular application are described in Appendix B. A recursive procedure is employed for parameter estimation; this has the advantage that all the data need not be stored in the computer for processing, but can be read from file one observation at a time, thus requiring very little computer storage. Data sequences with gaps also present no problems and can be processed in one operation, giving one set of estimated parameters.

The Constrained Linear Systems (CLS) approach hypothesizes that the response of a hydrological system to one or more inputs can be written as

$$q_t = \sum_{i=0}^{n_1} v_i^{(1)} u_{t-i}^{(1)} + \sum_{i=0}^{n_2} v_i^{(2)} u_{t-i}^{(2)} + \cdots + \sum_{i=0}^{n_r} v_i^{(r)} u_{t-i}^{(r)} + \varepsilon_t \tag{5.9}$$

where $v_i^{(1)}$, $v_i^{(2)}$, ... $v_i^{(r)}$ denote the ordinates of the impulse responses corresponding to the inputs $\{u_t^{(1)}, u_t^{(2)}, \ldots. u_t^{(r)}\}$, n_1, n_2, ... n_r.

represent the numbers of ordinates in the impulse responses, and ϵ_t is a noise term. The inputs $\{u_t^{(1)}, u_t^{(2)} \ldots u_t^{(r)}\}$ can be either upstream tributary flows or precipitation inputs; in estimating the ordinates of the corresponding impulse responses, constraints are imposed as follows in accordance with physical hydrological principles:

(a) inequality constraints : $v_i^{(1)}, v_i^{(2)}, \ldots, v_i^{(r)} \geq 0$ (5.10)

(b) equality constraints $\sum_{i=1}^{n_j} v_i^{(j)} = c_j,\ j = 1,2, \ldots,r$ (5.11)

where c_j in the case of a precipitation input corresponds to the observed coefficient of runoff, and in the case of a tributary flow input is equal to one in accordance with continuity.

Further details of the CLS model and parameter estimation procedure are given in Appendix B; thresholds can also be applied to the precipitation inputs in a similar fashion to that described for TF models.

The simple conceptual models which were employed assume that the runoff volume in each time interval can be derived by applying a coefficient of runoff to the rainfall input; this coefficient of runoff may be taken as constant throughout the year, in which case

$$ro_t = c.p_t \tag{5.12}$$

or to vary with season as

$$c_t = \{(c_{max} + c_{min})/2\} + \{(c_{max} - c_{min})/2\} \text{ Cos } \frac{2\pi(t - \phi)}{365} \qquad (5.13)$$

where c_{max} and c_{min} are maximum and minimum coefficients of runoff occurring throughout the year, and ϕ is a phase shift in days. Runoff is then generated as

$$ro_t = c_t.p_t$$

The generated runoff volumes ro_t are then routed through a linear reservoir with impulse response given by

$$v(t) = \frac{1}{k} e^{-t/k} \qquad (5.14)$$

to give the flow rate at the end of a particular time interval, or the average flow over that interval, depending on the form of the observed data with which model output is to be compared. The parameters occurring in these models are estimated through non-linear optimization (Appendix B).

5.2 Model calibration

5.2.1 Statistics of model fit

All of the models applied to the Citarum River Basin can be written in the form of equation (5.1), and the statistics of the errors η_t can be used to assess the goodness of fit over calibration and test periods. The statistics calculated during the present studies were the following:

$$\bar{\eta} = \frac{1}{n} \sum (q_t - \mathit{q}_t) \qquad (5.15)$$

$$\text{S.D. } (\eta_t) = [\frac{1}{n-1} \sum (\eta_t - \bar{\eta})^2]^{0.5} \qquad (5.16)$$

$$R^2 = (F_o^2 - F^2)/F_o^2 \qquad (5.17)$$

where

$$F_0^2 = \sum (q_t - \bar{q})^2 \tag{5.18}$$

$$F^2 = \sum (q_t - \mathit{q}_t)^2 \tag{5.19}$$

and where q_t denotes observed discharge and q_t denotes a deterministic simulation of discharge from a model. A value of $R^2 = 1$ corresponds to a perfect fit; a value of $R^2 = 0$ indicates that the simulation from the model is no better than would be obtained from the use of the mean discharge over the period of calibration.

5.2.2 Results for data set SA1

As discussed in Section 4, calibration studies were carried out using 4 different sets of rainfall data; data set SA1 corresponds to the early calibration runs carried out with the COSSARR model.

(a) Transfer function models: The steps involved in the identification of the appropriate order (r, s, b) for a TF model are described in Appendix B; a linear TF(1,1,0) or TF(1,2,0) model was found to be appropriate for the Nanjung basin. A TF(1,1,0) model has two parameters and is written as

$$\mathit{q}_t = \delta_1 \mathit{q}_{t-1} + \omega_0 P_t.$$

From Table 5.1 it can be seen that a TF(1,2,0) model ($R^2 = 0.516$) does not give significantly better results than a TF(1,1,0) model ($R^2 = 0.515$).

A multiple-input transfer function model for Palumbon was estimated for which the inputs were Nanjung flow (in mm over the catchment) and Palumbon Local rainfall (mm); the numbers of terms in the TF model were $r = 1$, $s_1 = 2$ (Palumbon Local rainfall) and $s_2 = 1$ (Nanjung flow), with notation TF(1/2,0/1,0). A value of $R^2 = 0.847$ was obtained; the parameter values in Table 5.1 illustrate that the model assigned a very heavy weight to Nanjung flow and relatively little weight to Palumbon Local rainfall, whereas, in the observed data, the proportion of flow at Palumbon due to Nanjung is less than that from the Palumbon Local basin. While measured Nanjung flow is clearly the best predictor of Palumbon flow, a model in which continuity is maintained for Nanjung flow would be preferable

Table 5.1 Statistics of model fits obtained using data set SA1

Catchment	Model	Inputs	Parameters	Error statistics		
				η	S.D. η_t	R^2
Nanjung	Linear TF(1,1,0)	Nanjung basin rainfall	$\delta_1 = -0.878$ $\omega_0 = 0.839$	0.483	46.22	0.515
	" TF(1,2,0)	Nanjung basin rainfall	$\delta_1 = -0.876$ $\omega_0 = 0.839$ $\omega_1 = 0.028$	0.464	46.18	0.516
Nanjung	SCM (1)	Nanjung basin rainfall	$c = 0.505$ $k = 8.373$ (days)	1.743	45.434	0.532
	SCM (2)	Nanjung basin rainfall	$c_{min} = 0.380$ $c_{max} = 0.580$ $k = 7.35$ days $\phi = 62^*$days	3.041	43.948	0.560
	SCM (3)	Nanjung basin rainfall	$c_{min} = 0.37$ $c_{max} = 0.59$ $k = 7.44$ days $\phi = 56$ days	3.196	43.917	0.560
Palumbon	Linear TF(1/2,0/1,0)	(1) Palumbon Local basin rainfall (2) Nanjung flow	$\delta_1 = -0.092$ $\omega_0^{(1)} = 2.450$ $\omega_1^{(1)} = 1.836$ $\omega_0^{(2)} = 30.073$	- 0.001	50.392	0.847

* denotes fixed parameter

on physical grounds; in addition, if a complete catchment simulation were to be carried out, the above model would be heavily dependent on the accuracy of the simulation of Nanjung flow which might not be desirable if a better simulation of the local Palumbon flow contribution could be obtained. Hence, this type of model for Palumbon was not pursued further.

(b) Simple conceptual models: Three versions of the models described in Section 5.1 were implemented :

Model SCM(1) : Parameters c, k

Model SCM(2) : Parameters c_{min}, c_{max}, k : ϕ(fixed) = 62 days

Model SCM(3) : Parameters c_{min}, c_{max}, k, ϕ

The value of ϕ for model SCM(2) was chosen on the basis that the soil moisture deficit in the Citarum River Basin is thought to be at a maximum around Aug.31. The results obtained for the three models are given in Table 5.1 for Nanjung; the value of R^2 = 0.53 obtained for model SCM(1) is very similar to that obtained for the TF(1,1,0) model for Nanjung as expected, since the same assumptions (a constant coefficient of runoff and the discrete time equivalent of the linear reservoir) are implicit in the TF(1,1,0) model. The improvements in R^2 for models SCM(2) (R^2 = 0.56) and SCM(3) (R^2 = 0.56) are not very significant, and suggest that the coefficient of runoff does not appear to vary significantly with season for the Nanjung basin. No runs were carried out with the CLS model for this data set.

5.2.3 Results for data set SA2

(a) Transfer function models

A linear TF(1,2,0) model was estimated for Nanjung with a value of R^2 = 0.603 (Table 5.2);the improvement in R^2 over data set SA1 (R^2 = 0.515) is partially attributable to the improved estimate of basin rainfall for data set SA2, and partially due to the elimination of an error in the program for computing basin rainfall which affected the results for data set SA1.

Table 5.2 Statistics of model fits obtained using data set SA2

Catchment	Model	Inputs	Parameters		Error	Statistics
				$\bar{\eta}$	S.D. η_t	R^2
Nanjung	Linear TF(1,2,0)	Nanjung basin rainfall	$\delta_1 = -0.810$ $\omega_0 = 1.250$ $\omega_1 = 0.881$	5.846	41.817	0.603
	TF(1,20) with threshold (k = 0.95; T = 220)	Nanjung basin rainfall (split)	$\delta_1 = -0.822$ $\omega_0^{(1)} = 0.716$ $\omega_1^{(1)} = 0.698$ $\omega_0^{(2)} = 1.308$ $\omega_1^{(2)} = 0.902$	2.937	40.213	0.633
	COSSARR	Nanjung basin rainfall		10.939	38.343	0.640
Palumbon	Linear TF(1,2,0)	(1) Nanjung basin rainfall (2) Palumbon basin rainfall	$\delta_1 = -0.667$ $\omega_0^{(1)} = 1.881$ $\omega_1^{(1)} = 0.590$ $\omega_0^{(2)} = 1.555$ $\omega_1^{(2)} = 3.480$	- 0.092	74.975	0.663
	COSSARR	(1) Palumbon local basin rainfall (2) Simulated Nanjung flow		2.889	64.885	0.747

A threshold was applied to an API computed using a constant coefficient $\bar{K}$ = 0.95 as described in Section 5.1; results for different values of the threshold are given in Table 5.2 illustrating that, for the best model (R^2 = 0.633), the improvement over the linear model is small.

A linear multiple input TF model was estimated for Palumbon, the inputs being Nanjung basin rainfall and Palumbon Local basin rainfall; a value of R^2 = 0.663 was obtained (Table 5.2).

(b) COSSARR model: The final calibration runs carried out with the COSSARRmodel used data set SA2; the statistics of model fit (Table 5.2) give R^2 = 0.64 for Nanjung and R^2 = 0.75 for Palumbon; for the latter model, the simulated flow at Nanjung was routed to Palumbon.

The result obtained for the best TF model with a threshold for Nanjung (R^2 = 0.63) is virtually identical to the fit obtained with the COSSARR model; for Palumbon, a TF model comparable to the COSSARR, with simulated Nanjung flow routed to Palumbon, was not developed for this data set. The result for the linear TF model with Nanjung and Palumbon Local basin rainfall inputs (R^2 = 0.66) compares favourably with the result for the SSARR model, given that the former model is linear and includes no routing component.

5.2.4 Results for data set NS1

(a) CLS models

A linear CLS model was calibrated for Nanjung with a value of R^2 = 0.72. For Palumbon, a CLS model was estimated in which Palumbon flow was related to upstream Nanjung flow and Palumbon Local basin rainfall; both equality and inequality constraints (5.10) and (5.11) were used in the estimation, thus ensuring that continuity was maintained for routed Nanjung flows. The value of R^2 obtained for this model was 0.86. The impulse response for Nanjung consisted of one ordinate of unit at lag zero, thus implying that average daily discharge for Nanjung can be translated directly to Palumbon. Using an API, a threshold was applied to the rainfall input for Palumbon Local but no significant improvement over R^2 = 0.86 was obtained. However, the number of runs which could be carried out with the CLS model was restricted since the program could only be executed when other programs were not being run on the Honeywell Mini.

(b) Transfer function models: A linear TF(1,1,0) was estimated for Nanjung, with $R^2 = 0.714$; this represents a significant improvement over the result for data set ($R^2 = 0.603$), and demonstrates the necessity of having an adequate estimate of basin rainfall for model calibration on the Citarum River. A TF(1,2,0) model did not give any improvement over the TF(1,1,0) and so a (1,1,0) structure was adapted for all further calibrations for Nanjung. TF models with thresholds were estimated using an API with a seasonally varying coefficient (equation 5.5) computed with $\bar{K} = 0.80$, $\alpha = 0.15$ and $\phi = 62$ days. The best model ($R^2 = 0.756$) was obtained with T = 90 (Table 5.3).

The result obtained above with the CLS model for Palumbon implied that a model for Palumbon Local flow could be developed separately by relating the difference between total Palumbon flow and Nanjung flow to Palumbon Local basin rainfall. A linear TF(1,1,0) model gave a value of $R^2 = 0.509$ which is much lower than the value of R^2 obtained for the linear TF(1,1,0) model for Nanjung (0.714). Inspection of the simulation errors suggested than an insufficient number of raingauges had recorded localized storms in a number of cases, and so the number of gauges was increased from 7 to 10; the results are presented in Section 5.2.5 (Data set NS2).

A multiple input linear TF model of total Palumbon flow (inputs Nanjung rainfall, Palumbon Local rainfall) gave a value of $R^2 = 0.675$, which represents only a marginal improvement over the result obtained for the same model with data set SA2 ($R^2 = 0.663$); although the Palumbon local basin produces the dominant contribution to Palumbon total flow, the improvement in the estimate of basin rainfall for Nanjung might have been expected to produce a better fit. To analyse this result further, a quantity called the gain can be computed for the inputs to the TF model; for data set SA2, the gains for the TF(1,1,0) model are

$$G_{NJ} = \frac{\omega_{0}^{(1)}.f}{1 - \delta_1} = 0.278$$

$$G_{PLL} = \frac{\omega_{0}^{(2)}.f}{1 - \delta_1} = 0.295$$

where f is a factor to take account of the different measurement units for rainfall and discharge. For data set NS1, the corresponding results are

Table 5.3 Statistics of model fits obtained using data set NS1

Catchment	Model	Inputs	Parameters	Error Statistics		
				$\bar{n}$	S.D. n_t	R^2
Nanjung	CLS (linear)	Nanjung basin rainfall	Impulse response ordinates	0.000	34.824	0.724
	Linear TF(1,1,0)	Nanjung basin rainfall	δ_1 = - 0.764 ω_0 = 2.465	2.343	35.496	0.714
	TF(1,1,0) with threshold $\bar{K}$ = 0.80 α = 0.15 ϕ = 62 T = 90.	Nanjung basin rainfall (split)	δ_1 = - .759 $\omega_0^{(1)}$ = 2.933 $\omega_0^{(2)}$ = 2.267	3.389	32.830	0.756
	SCM (1)	Nanjung basin rainfall	c = 0.530 k = 3.743	1.558	36.823	0.692
	SCM (2)	Nanjung basin rainfall	c_{min} = 0.376 c_{max} = 0.640 k = 3.470 ϕ = 62	2.357	33.356	0.746
	SCM (3)	Nanjung basin rainfall	c_{min} = 0.374 c_{max} = 0.639 k = 3.462 ϕ = 60.44	2.395	33.350	0.746
Palumbon Local	Linear TF(1,1,0)	Palumbon Local basin rainfall	δ_1 = - 0.741 ω_0 = 4.020	5.231	53.151	0.509
Palumbon (total area)	CLS (linear)	(1) Nanjung flow (measured) (2) Palumbon Local basin rainfall	Impulse response ordinates	0.000	50.620	0.859
	Linear TF(1,1,0)	(1) Nanjung rainfall (2) Palumbon Local rainfall	δ_1 = - 0.706 $\omega_0^{(1)}$ = 3.598 $\omega_0^{(2)}$ = 3.866	8.620	73.566	0.675
	Linear TF(1,1,0) Models for Nanjung and Palumbon Local combined	as above	as above	6.195	69.292	0.710
	SCM(1)	Palumbon basin rainfall	c = 0.489 k = 4.028	15.556	82.218	0.579
	SCM(2)	Palumbon basin rainfall	c_{min} = 0.361 c_{max} = 0.564 k = 3.893 ϕ = 62	16.834	77.159	0.625
	SCM(3)	Palumbon basin rainfall	c_{min} = 0.389 c_{max} = 0.585 k = 3.911 ϕ = 104.85	15.037	75.603	0.643

$$G_{NJ} = \frac{\omega_0^{(1)} f}{1 - \delta_1} = 0.260$$

$$G_{PLL} = \frac{\omega_0^{(2)} . f}{1 - \delta_1} = 0.279$$

Thus, for data set SA2, the model predicts that 0.278 and 0.295 of Nanjung and Palumbon Local basin rainfalls, respectively, will become flow, comparison with the results for data set NS1 shows that, while the absolute values of the gains have changed slightly, their ratio has not and so the model has not assigned more weight to Nanjung rainfall. This is consistent with the improvement obtained in R^2, but it is a little surprising that the model did not assign more weight to Nanjung rainfall.

The simulated flows obtained from the linear TF(1,1,0) model for local Palumbon flow (R^2 = 0.509) were added to the simulated flows for the linear TF(1,1,0) model for Nanjung (R^2 = 0.714) to obtain total simulated flow at Palumbon, with a calculated value of R^2 = 0.710. Hence, this approach provides a better simulation of total Palumbon flow than the multiple input model discussed above.

(c) Simple conceptual models : Models SCM(1)-(3) were calibrated for Nanjung and total Palumbon flows; in the latter case, all of the stations used to estimate basin rainfall for the Nanjung and Palumbon Local areas were used to provide a single estimate of total basin rainfall at Palumbon (area 4061 km^2). The results in Table 5.3 show that, for Nanjung, the SCM(1) model gives a similar R^2 value (0.692) to that for the TF(1,1,0) model (0.714) as expected; an improvement to R^2 = 0.746 was obtained with the seasonally varying coefficient of runoff. For Palumbon, the corresponding results are R^2 = 0.580 and R^2 = 0.643; thus, the use of a routing component for Nanjung to Palumbon gives a better model for Palumbon (R^2 = 0.710) than a total basin rainfall-runoff model.

5.2.5 Results for data set NS2

As noted in Section 4, the number of stations used in calculating Palumbon Local basin rainfall was increased from 7 to 10 for this data set; the number for Nanjung remained unchanged. A linear TF(1,1,0) model for Palumbon local flow gave R^2 = 0.567, compared with R^2 = 0.509 for data set NS1; this result suggests that 7 stations, and perhaps 10,

Table 5.4 Statistics of model fits obtained using data set NS2

Catchment	Model	Inputs	Parameters	Error Statistics		
				$\bar{\eta}$	S.D. η_t	R^2
Palumbon Local	Linear TF(1,1,0)	Palumbon Local basin rainfall	$\delta_1 = -0.693$ $\omega_o = 4.974$	4.549	49.936	0.567
	TF(1,1,0) with threshold $\bar{K} = 0.80$ $\alpha = 0.15$ $\phi = 62$ $T = 60$	Palumbon local basin rainfall	$\delta_1 =$ $\omega_o^{(1)} =$ $\omega_o^{(2)}$			0.581
Palumbon (Total area)	Linear TF(1,1,0)	(1) Nanjung basin rainfall (2) Palumbon basin rainfall	$\delta_1 = -0.685$ $\omega_o^{(1)} = 3.270$ $\omega_o^{(2)} = 5.084$	6.578	70.392	0.703
	TF(1,1,0) with Threshold for Nanjung and Palumbon Local basins	Nanjung and Palumbon Local basin rainfalls with simulated Nanjung flow routed to Palumbon	as for TF(1,1,0) model with T = 90 for Nanjung (Table 5.3) and for TF(1,1,0) model with T = 60 for Palumbon Local basin given above	7.302	95.690	0.745

are insufficient to provide an accurate estimate of Palumbon Local basin rainfall. The best result for a TF(1,1,0) model with a threshold was R^2 = 0.581 (T = 60), while for the multiple input model (Nanjung rainfall, Palumbon Local rainfall), a value of R^2 = 0.703 was obtained, compared with R^2 = 0.675 for data set NS1. By combining the best models for Nanjung (TF(1,1,0) with T = 90) and Palumbon Local flows (TF(1,1,0) with T = 60) a value of R^2 = 0·745 resulted for total simulated flow at Palumbon (Table 5.4).

5.3 DISCUSSION

At the time of writing, the results from the COSSARR model calibrations on data set NS1 were not available, and so a direct comparison with the results from the simple models can only be made for data set SA2. For Nanjung, a TF(1,1,0) model with a threshold gave a similar fit to the data (R^2 = 0.63) as the COSSARR model (R^2 = 0.64); however, for Palumbon, the latter model (R^2 = 0.75) gives a better result than the linear TF(1,2,0) model with Nanjung and Palumbon Local basin rainfall inputs (R^2 = 0.67). A better non linear TF model for Palumbon was not sought with this data set, since it was noted that, with data set NS1, large errors in simulated discharge were attributable to errors in sampling localized rain storms. The progressive improvement in the results from the simple models for data sets SA1, SA2, NS1 and NS2 is almost entirely attributable to improvements in estimating the basin rainfall inputs; hence, it is surprising that the COSSARR model achieved such a good result for Palumbon with data set SA2.

The best TF model (with threshold) obtained for Nanjung gave R^2 = 0.746 for Nanjung (data set NS1) while the best result which could be obtained for Palumbon Local flow was R^2 = 0.581 (data set NS2). The disparity in these R^2 values is surprising; however, an improved fit was obtained for Palumbon Local flow when the number of raingauges used in computing basin rainfal was increased (7 for data set NS1 to 10 for NS2); a further improvement in fit might be obtained by increasing the number of gauges still further. The raingauge densities for Nanjung for data set NS1 (1 per 245 km^2) and Palumbon Local for data set NS2 (1 per 234 km^2) are very similar; a closer analysis of the sampling of rain storms over the Palumbon Local area would be required to establish if this can account for the difference in fit obtained for the two catchment areas.

The best overall model for Palumbon was obtained by combining the best TF models for these two catchment areas to give a value of R^2 = 0.75 i.e. the same as that obtained for the COSSARR model with data set SA2. However, the COSSARR model might be expected to do better with data set NS1.

6. EVALUATION OF RESULTS FOR 'TRIAL RUN' FORECAST PERIOD, DECEMBER 1980 - APRIL 1981.

6. EVALUATION OF RESULTS FOR 'TRIAL RUN' FORECAST PERIOD DECEMBER 1980 - APRIL 1981

6.1 Rainfall forecasting

The procedure used for making one and two-day ahead rainfall forecasts for the Citarum River Basin has been described in outline in Section 3; to assess how this procedure performed over the operational test period December 1980-April 1981, the one day and two day forecasts were punched up together with the observed values of average rainfall for the Nanjung and Palumbon Local basins, and the mean, standard deviation and R^2 value (5.15-5.17) of the forecast errors were computed for those days for which forecasts were made. The mean and standard deviation of observed rainfall are also presented in Table 6.1; the calculated values of R^2 indicate that the forecasting procedure gives slightly better results than the use of the mean $\bar{R}$ of the set of observations as the forecast which corresponds to $R^2 = 0$. However, the mean $\bar{R}$ would not be known a priori and so this does not constitute a basis for an operational comparison with the Project procedure. The following alternative procedures were employed to provide a basis for assessing the Project procedure:

(a) a procedure which specifies that the rainfall on days $(t + 1)$ and $(t+2)$ will be the same as on day t ;

(b) use of an autoregressive moving average (ARMA) time series model with parameters estimated from historical rainfall data i.e.

$$R_t = \bar{R} - \phi_1(R_{t-1} - \bar{R}) - \phi_2(R_{t-2} - \bar{R}) - \ldots - \phi_p(R_{t-p} - \bar{R})$$

$$+ \; a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q} \qquad (6.1)$$

where $\bar{R}$ is average rainfall, $\phi_1 \ldots \phi_p$ are p autoregressive parameters, $\theta_1, \ldots \theta_q$ are q moving average parameters and a_t is an independently distributed random variable with zero mean.

(c) use of the recursively estimated mean of the observations over the forecast period i.e. the mean of the set of observations up to the current time point is used as the forecast :

Table 6.1 Statistics of 1-day and 2-day ahead Project rainfall forecast errors for the Nanjung and Palumbon Local basins for the period December 1, 1980 - April 30, 1981

	Nanjung Basin		Palumbon Local Basin	
	1 day	2 day	1 day	2 day
n	119	117	119	117
$\bar{R}$	5.488	5.791	7.249	7.543
S.D.(R_t)	5.916	5.888	7.754	7.794
mean error	- 0.030	0.466	0.680	1.312
st. dev.	5.810	5.677	7.351	7.525
R^2	0.043	0.072	0.101	0.048

Table 6.2 Fitted parameter values and R^2 statistic for AR(4) and AR(3) models fitted to Nanjung and Palumbon Local Basin rainfalls, respectively, for data set NS1.

(a) Nanjung Basin

Model	Parameters					R^2
AR(4)	$\bar{R}$	ϕ_1	ϕ_2	ϕ_3	ϕ_4	
all data	7.219	-0.422	-0.010	--0.010	-0.143	0.288
w.s. data	9.359	-0.395	-0.104	-0.027	-0.087	0.217

(b) Palumbon Local Basin

Model	Parameters				R^2
AR(3)	$\bar{R}$	ϕ_1	ϕ_2	ϕ_3	
all data	6.216	-0.419	-0.043	-0.103	0.234
w.s. data	7.806	-0.434	-0.036	-0.046	0.207

$$\bar{R}_t = \bar{R}_{t-1} + \frac{1}{t}(R_t - \bar{R}_{t-1}) \tag{6.2}$$

Historical rainfall data for the period 1974-76 were used to identify and fit ARMA(p,q) models for Nanjung and Palumbon Local average basin rainfalls; results for data set NS1 are presented here. Models were developed using (i) all the daily data within each year and (ii) using only wet season daily data (November 1 - April 30). AR(4) and AR(3) models were identified for Nanjung and Palumbon Local Basin rainfalls, respectively (both for all data and wet season data); the estimated parameters and values of R^2 are given in Table 6.2.

The results obtained when procedures (a)-(c) were applied to the data for the period December 1, 1980 to April 30, 1981 are presented in Table 6.3; for the autoregressive models, the parameter values used were those given in Table 6.2 for 'all data', 1974-76. As expected the $R_{t+2} = R_{t+1} = R_t$ model performs worst but serves as a baseline for comparison; the remaining procedures give results in the neighbourhood of $R^2 = 0$. The results for the Project forecasts presented in Table 6.1 are somewhat better than the best results in Table 6.3, although not by a significant amount. The results obtained for the AR models could probably be improved by applying these models to longer series of data; also, the use of recursive parameter estimation in real-time for such models might also lead to improved results.

6.2 Flow forecasting

As described in Section 3, forecasts of discharge at 07.00 hours at Nanjung and Palumbon one and two days ahead were made during the 'trial run' test period December 1, 1980 - April 30, 1981; from these, and forecasts of the Jatiluhur Local flow contributions, forecasts of Jatiluhur reservoir level were computed. The statistics of the one day and two day ahead forecasts of Nanjung discharge, Palumbon discharge and Jatiluhur reservoir level at 07.00 hours are given in Table 6.4; with the exception of Jatiluhur reservoir levels, the values of R^2 are relatively low, since they are heavily influenced by the errors in the rainfall forecasts.

Table 6.3 Statistics of 1-day and 2-day ahead rainfall forecast errors for various procedures applied to the Nanjung and Palumbon Local basins

(a) Nanjung Basin

	$R_{t+2} = R_{t+1} = R_t$		AR(4)		Recursively Est. Mean	
	1 day	2 day	1 day	2 day	1 day	2 day
n	119	117	119	117	119	117
mean	0.018	0.019	-0.762	-1.055	-1.125	-1.141
st. dev.	7.295	7.800	5.830	5.831	5.948	5.98
R^2	-0.508	-0.724	0.021	0.005	-0.039	-0.051

(b) Palumbon Local Basin

	$R_{t+2} = R_{t+1} = R_t$		AR(3)		Recursively Est. Mean	
	1 day	2 day	1 day	2 day	1 day	2 day
n	119	117	119	117	119	117
mean	-0.186	0.297	0.400	0.748	-1.170	-1.099
st. dev.	10.781	10.146	8.328	7.817	8.018	8.081
R^2	-0.906	-0.690	-0.140	-0.011	-0.077	-0.091

Table 6.4 Statistics of one day and two day ahead COSSARR forecast errors for Nanjung flow, Palumbon flow and Jatiluhur reservoir level at 07.00 hours. The mean and standard deviation of observed discharge for Nanjung were 85.1 and 56.9 and for Palumbon were 238.0 and 145.2 m^3/s, respectively.

Location	Nanjung		Palumbon		Jatiluhur	
Lead time	1 day	2 day	1 day	2 day	1 day	2 day
n	112	112	112	112	109	109
mean	9.286	14.868	13.712	18.413	-0.020	0.031
st.dev.	37.469	44.977	120.047	104.608	0.207	0.255
R^2	0.543	0.323	0.314	0.354	0.996	0.994

The 1 day ahead forecasts for Palumbon (R^2 = 0.314) are better than the two day ahead forecasts (R^2 = 0.354) which is surprising, given that the reverse is true for the rainfall forecasts (Table 6.1); - also the one day ahead forecasts for Nanjung (R^2 = 0.543) are better than those for Palumbon (R^2 = 0.314). As the simulation model results for Palumbon (R^2 =0.75) were better than for Nanjung (R^2 = 0.64) this result may reflect the updating procedure used with the COSSARR model.

The very high values of R^2 observed for Jatiluhur reservoir levels reflect the fact that large errors in forecasted inflows translate to small errors in reservoir level forecasts because of the large surface area of the reservoir. This raises the question as to what the desired accuracy in forecasting reservoir levels should be. It is when high rainfalls occur that forecasts of reservoir level are likely to be of greatest importance; however, as the Project rainfall forecasting procedure underestimates considerably the magnitude of high rainfall over the Citarum Basin, forecasts of high discharge tend to be made one day late i.e. after the rainfall has been observed and measured discharge is already high, thus detracting from the value of the forecasts.

To allow the simple TF models developed in Section 5 to be used for real-time forecasting, a procedure for updating model forecasts in real-time is required; this is achieved by developing a noise model for the term η_t i.e the difference between observed flow q_t and the simulation obtained from the TF model q_t. An ARMA(p,q) model can be used to describe the structure of the η_t, and a recursive procedure applied to estimate the model parameters (Appendix B) from the η_t series derived from fitting the TF model over the calibration period; the composite model is called a transfer function noise (TFN) model. Noise models were estimated for the η_t series obtained from fitting linear TF(1,1,0) models at Nanjung and Palumbon for data set NS1 (Table 5.3); AR(4) models were found to be appropriate in each case, and the parameter values and R^2 values are given in Table 6.5.

Table 6.5. Parameters and R^2 values for AR(4) models estimated from η_t series for linear TF(1,1,0) models fitted using data set NS1

Nanjung					Palumbon				
ϕ_1	ϕ_2	ϕ_3	ϕ_4	R^2	ϕ_1	ϕ_2	ϕ_3	ϕ_4	R^2
-0.672	-0.107	0.053	-0.118	0.633	-0.491	-0.108	-0.076	-0.103	.451

The values of R^2 indicate that there is more persistence in the η_t series for Nanjung than for Palumbon.

The TFN models were applied in simulated 'real-time' mode over the test period December 1980 - April 1981; as forecasts of discharge at 07.00 hours one and two days ahead had been provided by the COSSARR model, similar forecasts were required from the TFN model for comparison purposes. However, the TFN models had been calibrated using average daily discharge data, and thus should be used to forecast these quantities. This was not possible since only three discharge values (at 0700, 1100 and 1700 hours) were recorded during the test period, and only the value at 07.00 hours was available at the time the COSSARR forecasts for the next two days were made. Hence, the TFN model was used to provide forecasts of discharge at 07.00 hours although such values would not be expected to be representative of average daily discharge which the model had been calibrated on. The mean,standard deviation and R^2 values for the one day and two day ahead TFN forecasts are given in Table 6.6; comparison with Table 6.4 shows that the results are somewhat better overall than those for the COSSARR model. The two day ahead forecasts for Palumbon are slightly better than the one day forecasts; this result was also obtained with the COSSARR model (Table 6.5). In producing the forecasts from the TFN model, the observed rainfall used up to the 'current' time point was that computed when the data from all the reporting stations in the Citarum Basin had been received. This appears to give an advantage to the TFN model since the 'observed rainfall' used by the COSSARR model was based largely on the data obtained from the network of stations reporting in real-time to the Project Office. However, it is unlikely that this would make much difference to the results, since, in real-time, observed discharge data are available up to the current time point, and the noise model can compensate for any inadequacies in the simulation from the TF model due to the rainfall input.

Table 6.6 Statistics of 1 day and 2 day ahead forecast errors at 07.00 hours for Nanjung and Palumbon for the TFN model using forecasted rainfall

	Nanjung		Palumbon	
	One day	Two day	One day	Two day
Mean	5.617	11.239	19.818	24.829
St.dev.	35.456	40.469	110.645	97.601
R^2	0.605	0.468	0.406	0.419

The linear TFN models have also been run assuming perfect knowledge of future rainfall i.e. observed rainfall is used instead of forecasted rainfall in making one day and two day ahead forecasts of flow. For example, the value of R^2 for one day ahead forecasts at Palumbon is 0.692 (Table 6.7) compared with 0.406 for forecasted rainfall (Table 6.6); *this illustrates the large component of error that is attributable to* forecasted rainfall. In Table 6.7, the value of R^2 for two day ahead forecasts is higher than for one day ahead. To explain this result, the forecast errors have been inspected and it has been found that a number of large one day ahead errors occurred on days for which no two day ahead Palumbon forecast was made, and so the apparent anomaly is attributable to the different sub-sets of forecasts used to calculate the statistics.

Table 6.7 Statistics of one day and two day ahead forecast errors at 07.00 hours for Nanjung and Palumbon for the TFN model using future observed rainfall

	Nanjung		Palumbon	
	One day	Two day	One day	Two day
Mean	5.485	9.394	15.717	14.374
St. dev.	26.391	25.733	79.459	58.140
R^2	0.777	0.774	0.692	0.795

It is also of interest to see how well the TF model performs when used in simulation mode to reconstitute the flows at Nanjung and Palumbon for the 'trial run' period; the results for the linear TF models used for 'real-time' forecasting over this period are given in Table 6.8. The values of R^2 obtained are much lower than for the fitting period (1974-76); this is largely due to a consistent underestimation of the flows at 07.00 hours. Since the peak daily flow rates for Nanjung and Palumbon usually occur in the early hours of the morning, flow at 07.00 hours will tend to be consistently higher than average daily flow, thus accounting for the large positive values of $\bar{\eta}$ in Table 6.8. If these biases are corrected for in computing the R^2 values i.e. F^2 in (5.19) is computed as

$$F^2 = \Sigma (\eta_t - \bar{\eta})^2$$

then the resulting values of R^2 are much higher (Table 6.8) and similar to the values obtained over the fitting period (Tables 5.3 and 5.4).

Table 6.8 Statistics of simulation errors when linear TF(1,1,0) models used to reconstitute flows over period December 1980 - April 1981 at Nanjung and Palumbon

	Nanjung	Palumbon
$\bar{\eta}$	31.154	70.034
s.d. η_t	33.922	84.741
R^2	0.3910	0.407
R^2(corrected)	0.6707	0.648

6.3 DISCUSSION

The results presented in Section 6.1 suggest that the procedure used within the Project for rainfall forecasting utilizes the available information effectively; however, the forecasts are statistical in nature, rather than deterministic, and so should have some error bounds or confidence limits quoted with them. These, when translated into forecasts of flow and reservoir level, would give the user an idea of the uncertainty associated with the forecasts.

Despite the fact that the simple TFN models had been calibrated using average daily discharge data, the TFN model forecasts of flow at 0700 hours for Nanjung and Palumbon were better than the COSSARR model forecasts over the 'trial-run' period. This result is largely due to the ability of the noise model to update forecasts efficiently in real-time, since the TF model gave a relatively poor simulation of discharge at 0700 hours which is largely attributable to the fact that the latter flow rate is consistently higher than average daily flow. Since it is likely that the COSSARR model would have provided a better reconstitution of flow than the TF model over the 'trial run' period, the COSSARR updating procedure (i.e. adjustment of observed rainfall over the 'back-up' period') is probably not as efficient as that used with the TFN model.

The results obtained when the TFN models were used to make forecasts assuming perfect knowledge of future rainfall illustrate the extent to which the errors in flow forecasts are dominated by errors in rainfall forecasts; under these conditions, there is little to be gained by using complex models since any improvement that might be obtained with a complex model over a simple model is liable to be small compared with that which could result from improved rainfall forecasts. The results presented above suggest that the simpler TFN model can perform as well as, if not better, than the more complex COSSARR model, and so under these circumstances, little appears to be gained from using the more complex model.

To enable the TFN model to be used to forecast Jatiluhur reservoir levels, forecasts of flow at 0700 hours at Palumbon would need to be converted to forecasts of average daily flow into the reservoir, and a reservoir routing component added to provide forecasted reservoir levels. There was insufficient time to allow this work to be undertaken within the consultant's assignment.

7. CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS AND RECOMMENDATIONS

The COSSARR model has, within the limits imposed by the available data, performed adequately as a forecasting tool over its first 'trial run' period; however, this conclusion must be qualified with the following considerations:

(i) in calibrating the COSSARR model, better results would have been obtained if stations with infilled data had not been used in estimating basin rainfall;

(ii) simple transfer-function models with few parameters have been calibrated for Nanjung and Palumbon during the consultant's visit and shown to give comparable results to those obtained with the COSSARR model;

(iii) the forecasts of daily rainfall one and two days ahead made within the Project are as good as can be obtained with the available information;

(iv) while the number of reporting stations used to estimate basin rainfall in real-time is relatively small, the importance of this is diminished by the adjustment of the rainfall input to the COSSARR model during the 'back-up period';

(v) the real-time forecasts obtained using the simple model developed under (ii) above over the 'trial run' period were slightly better than those obtained from the COSSARR model; this suggests that the COSSARR forecast updating procedure could be improved upon;

(vi) the full capability of the COSSARR model to simulate river regulation by a complex system of reservoirs is not required for the Citarum River Basin

Taking the foregoing conclusions into consideration, forecasts of similar accuracy to those produced by the COSSARR model can be made by simpler models. In cost/benefit terms, the level of benefits from both models would be the same but the implementation and running costs for the COSSARR model would be much higher : in the case of the Citarum River, 4-6 weeks consultant's time is estimated for simple model implementation, and 6 months for COSSARR. Once calibrated the simple model can be run on a cheap desk-top micro-computer while COSSARR, developed to handle a general configuration of rivers and reservoirs, requires a larger facility.

The following recommendations are made by the consultant:

(a) Effort should be devoted to improving the rainfall forecasts; this can only come if more quantitative meteorological information (e.g. from radar) is supplied in real-time by BMG;

(b) the simple models developed by the consultant should be run operationally alongside the COSSARR model during the 1981/82 'operational run' forecasting period;

(c) no effort should be made to implement further complex models (e.g. the Stanford Watershed Model) under the Project since

   (i) the data to support such models do not exist for the Citarum River Basin;

   (ii) even if sufficient data were available, the results presented in Section 6 show that the factor limiting the accuracy of flow and reservoir level forecasts is the accuracy of the rainfall forecasts;

(d) if the COSSARR model were to be transferred to other river basins, then

   (i) appropriate computing facilities would be needed to run the model at the various forecasting centres;

   (ii) staff would have to be trained in its use.

Hence, considerable resources would be required for implementation on a multi-basin scale. Before any such transfers are contemplated, the following steps should be taken:

A. the benefits should be carefully assessed, both in terms of transfer of knowledge and for operational flood warning, reservoir management etc.

B. calibration studies should be carried out at DPMA with simple models and the COSSARR model, and the models then run over hypothetical trial forecasting periods; unless the COSSARR can be shown to give significantly better results, the simple models should be adapted for implementation;

the development of a forecasting capability for other river basins throughout Indonesia would be best undertaken through the operation of a forecasting model development centre, with some specialist support, at DPMA; simple models could then be transferred to regional centres to be run operationally on minimum cost desk-top computers where required throughout Indonesia. In this way the expertise and computing facilities concentrated at DPMA would be exploited to maximum effect.

APPENDIX A: SUMMARIES OF LECTURES

RAINGAUGE NETWORK RATIONALIZATION

Summary

In the United Kingdom, there are about 6,500 daily, weekly or monthly-read raingauges; the collection, processing and dissemination of the rainfall data are shared by the Meteorological Office and the ten Regional Water Authorities who are responsible for all aspects of water resources planning and management within their respective areas. The costs of collecting and processing rainfall data have increased in recent years, and reservations have been expressed about the quantity of data which are collected and processed. To establish whether the UK raingauge network fulfils its role in the most cost effective way, a project was undertaken jointly by the Institute of Hydrology and the Meteorological Office to develop methods of evaluating raingauge networks and for redesigning them, and to apply these techniques to some of the Regional Water Authority networks in the UK.

A network of raingauges provides information about rainfall at only a limited number of points within a region; some procedure must then be adopted to estimate the rainfall for other chosen points and areas within the region; in addition, the accuracy of rainfall estimates must be quantified so that a comparison with the requirements of user of rainfall data can be made. Optimal estimation procedures have been developed which minimize the mean square error of estimation; these can be applied to areas with any number and configuration of raingauges. The techniques can be applied to the redesign of existing networks of gauges by mapping the root mean square error of point interpolation, allowing the identification of these areas where these are surplus gauges or where new gauges are needed to meet some specified criterion of accuracy.

The techniques which have been developed have been applied to the re-design of the Wessex Water Authority raingauge network in Southern England.

RAINFALL - RUNOFF MODELLING

Summary

Over the past twenty years, considerable research effort has been devoted to the development of mathematical models of the rainfall-runoff process. While the main scientific objective of such work has been to obtain a better understanding of the complexity of catchment response, the reseach has also been motivated by the necessity for such models in the short-term management of water resources. One of the main potential areas of application of rainfall-runoff models is in the short-term forecasting of streamflow, where forecasts from such models form the basis of decisions pertaining to flood warning, flood control or river regulation.

Given the considerable number of rainfall-runoff models which have been developed to date, the question then arises as to what type of model is most suitable for real-time use. In this context, it is useful to distinguish between three types of model.

(a). Distributed physics-based models:

With such models, the objective is to use the equations of mass, energy and momentum to describe the movement of water over the land surface and through the unsaturated and saturated zones. The resulting system of partial differential equations has to be solved numerically at all points on a three dimensional grid representation of a catchment system. Such models are very much at the development stage at present (eg. the European Hydrological System, Jonch-Clausen, 1979) but will eventually offer the possibilities of satisfactorily predicting the hydrological effects of and use changes, and of satisfactorily predicting the response of ungauged catchments.

(b). Lumped conceptual models:

The essence of these models is that they are quasi physical in nature; rather than using the relevant equations of mass, energy and momentum to describe the component processes of the rainfall-runoff process, simplified but plausible conceptual representations of these processes are adopted. These representations frequently involve several interlinked

stores and simple budgeting procedures which ensure that at all times a complete mass balance is maintained between all inputs, outputs and inner storage changes. The forerunner of this type of conceptual model is the Stanford Watershed Model developed originally by Crawford and Linsley (1963).

(c). <u>Input - output or black box models</u>:

With such models, attention centres on identifying a relationship between rainfall input and streamflow output without attempting to describe the internal mechanisms whereby this transformation takes place. This approach is frequently referred to as the systems approach, as it relies heavily on techniques of systems analysis. A classical example of a model of this type is the unit hydrograph which postulates a linear relationship between 'effective rainfall' and 'storm runoff' and which can be identified using any one of a number of techniques of input - output analysis.

Examples of each of the above types of model will be given and their suitability for real-time flow forecasting will be discussed.

REAL-TIME FLOW FORECASTING

Summary

With the increasing use of telemetry in the control of water resource systems, a considerable amount of effort is being devoted to the development of models and parameter estimation techniques for real-time use. Of particular importance is the use of efficient computational procedures for updating flow forecasts as new data are received in real-time.

The various procedures which can be used for updating flows forecasts from conceptual models are discussed; these include adjusting the model parameters, adjusting the contents of the various storages, adjusting the rainfall input or employing a stochastic model to forecast the residuals obtained from the simulation model. In the case of input-output models, more sophisticated recursive estimation procedures can be employed which update model forecasts recursively in real-time. These procedures when used with simple input-output models require minimal computational facilities.

After outlining the basic principles of recursive estimation, a particular class of input-output models suitable for real-time use will be described; these models are called transfer function noise models, and they employ a recursive procedure for parameter estimation. The basic transfer function model is linear; procedures for introducing non-linearity into these models will be described. Some results obtained from applying them to some British catchments will be presented.

APPENDIX B: DESCRIPTIONS OF SIMPLE FLOW FORECASTING MODELS

B.1. TRANSFER FUNCTION NOISE MODELS

B.1 TRANSFER FUNCTION NOISE (TFN) MODELS

B.1.1 Transfer function (TF) models

The basic modelling approach assumes that the observed flow q_t can be considered to be the sum of a linearly deterministic component $\mathcal{q}_t$ and a stochastic component η_t :

$$q_t = \mathcal{q}_t + \eta_t . \qquad \text{(B1.1)}$$

The deterministic component of flow, $\mathcal{q}_t$, is defined such that it is exactly related to the input, u_t, (which may be rainfall p_t or some other quantity which influences flow) by the deterministic linear model

$$\mathcal{q}_t + \delta_1 \mathcal{q}_{t-1} + \dots + \delta_r \mathcal{q}_{t-r} = \omega_o u_{t-b} + \omega_1 u_{t-b-1} + \dots + \omega_{s-1} u_{t-b-s-1}$$

where δ_i, ω_i are parameters, and b is the pure time delay before the flow output responds to a change in the rainfall input. Introducing the backward difference operator, B, defined in $B^b u_t = u_{t-b}$, the above may be expressed concisely in difference equation form as:

$$\delta(B)\, \mathcal{q}_t = \omega(B)\, u_{t-b} \qquad \text{(B1.2)}$$

where the autoregressive and moving average operators $\delta(B)$ and $\omega(B)$ are polynomials in B of degree r and s-1 respectively, i.e.,

$$\delta(B) = 1 + \delta_1 B + \dots + \delta_r B^r,$$

$$\omega(B) = \omega_o + \omega_1 B + \quad + \omega_{s-1} B^{s-1}. \qquad \text{(B1.3)}$$

On writing (B1.2) as

$$\mathcal{q}_t \qquad \delta^{-1}(B)\, \omega(B)\, B^b u_t = v(B) u_t, \qquad \text{(B1.4)}$$

where

$$v(B) = v_o + v_1 B + v_2 B^2 +$$

the series of coefficients v_o, v_1, ... is called the system impulse response function. This is equivalent to the unit hydrograph encountered in the hydrological literature, except that the coefficients are used to define the relationship between <u>total</u> flow and rainfall and are not constrained to sum to unity; in fact their sum is called the gain of the system which may be equated to the runoff coefficient of a catchment. Since in general the order of a polynomial approximation to $v(B)$ will be larger than the sum of the orders of $\delta(B)$ and $\omega(B)$, the form of (B1.2) offers important advantages by virtue of its parametric efficiency (Box and Jenkins, 1970). Also the dependence of current flow on past flows (that is, the autoregressive nature of (B1.2)) is of particular importance in real-time forecasting applications: it provides a natural mechanism whereby forecasts may be based on the most recently observed values of flow, and not just on past rainfall as in the unit hydrograph (or impulse response function) representation (Moore and O'Connell, 1978).

B1.2 ARMA noise models

The stochastic component, η_t, is attributed to the aggregated disturbance effects of model errors, and measurement errors. It is normal practice when estimating the unit hydrograph (UH) ordinates to assume that the model errors, η_t, form an uncorrelated sequence: this assumption allows the UH ordinates to be estimated, for example by least squares (Snyder, 1955). However in general the noise, η_t, will not form an uncorrelated sequence, and inefficient parameter estimates will result if least squares is used. The noise, η_t, can be reasonably assumed to be related to white noise (an uncorrelated sequence of random variables) by the difference equation

$$\phi(B)\eta_t = \theta(B)a_t \qquad \text{(B1.5)}$$

where the autoregressive and moving average operators are defined as

$$\begin{aligned} \phi(B) &:= 1 + \phi_1 B + \quad + \phi_p B^p, \text{ and} \\ \theta(B) &:= 1 + \theta_1 B + \quad + \theta_q B^q \end{aligned} \qquad \text{(B1.6)}$$

respectively. The white noise sequence, a_t, is assumed to have zero mean and variance σ_a^2, and to be uncorrelated with the input, u_t.

The deterministic transfer function (TF) component of the model represented by equation (B1.2) will be referred to as the process model, and equation (B1.5) relating the process noise, η_t, to white noise, a_t, will be referred to as the noise model. Eliminating q_t, by combining (B1.2) and (B1.5) using (B1.1), gives the composite, or transfer function noise (TFN) model

$$q_t = \frac{\omega(B)}{\delta(B)} u_{t-b} + \frac{\theta(B)}{\phi(B)} a_t, \qquad \text{(B1.7)}$$

depicted in Figure B.1. This composite model thus not only provides a more efficient parameterisation than the UH representation of a linear system but also, by acknowledging that the hydrological system is stochastic, provides a model for the correlated model residuals which can be used to improve upon the deterministic forecast of flow provided by the process model.

When discussing different types of process model in later sections it will be found convenient to express the transfer function noise model (B1.7) in the form

$$\delta(B)q_t = \omega(B)u_{t-b} + \varepsilon_t \qquad \text{(B1.8)}$$

where

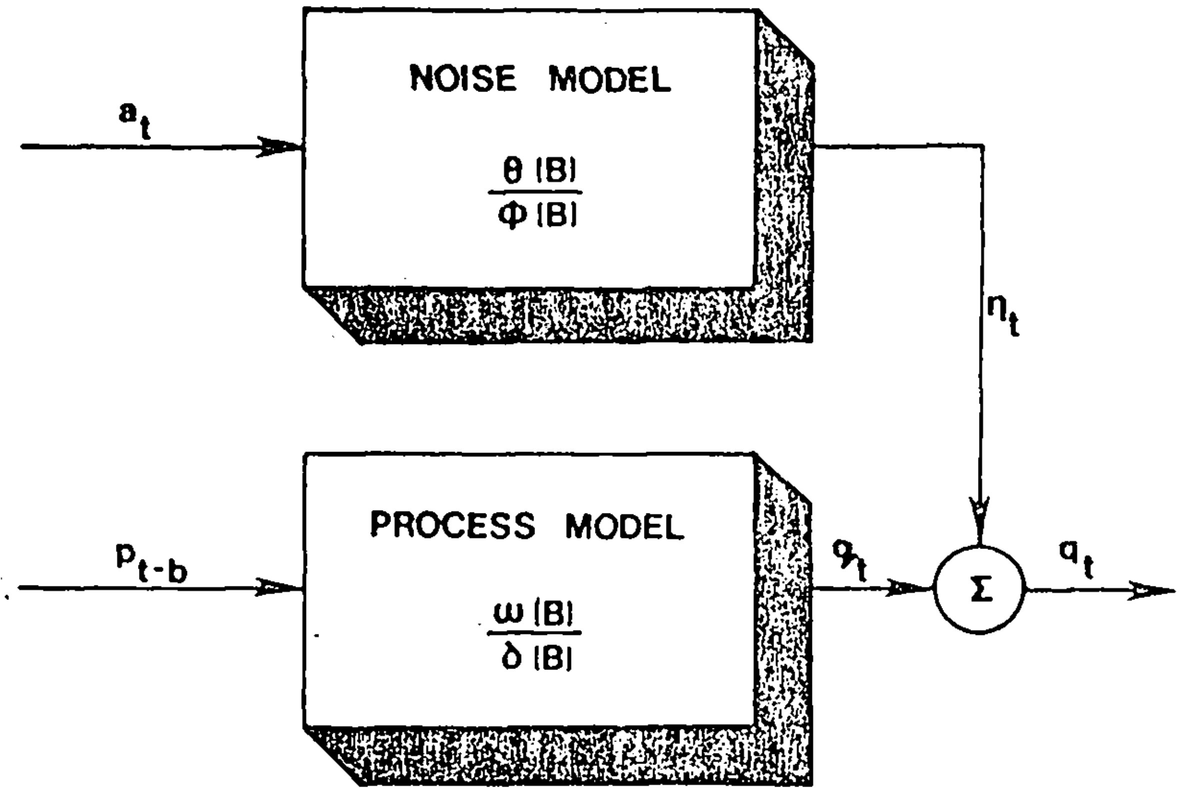


Figure B.1 Transfer function noise model

$$= \frac{\delta(B)\theta(B)}{\phi(B)} a_t = \delta(B)\eta_t$$

The particular structure of the transfer function model will be indicated by TF(r, s, b), where r and s specify the orders of the polynomials $\delta(B)$ and $\omega(B)$ in (B1.3) and b is the pure time delay; similarly the structure of the noise model will be indicated by the notation ARMA(p,q) where p and q are the orders of the polynomials $\phi(B)$, $\theta(B)$ defined in (B1.6).

B1.3 Extension to the multiple input case

The process model is readily extended to the case where several hydrological inputs are considered to influence flow:

$$\delta(B)q_t = \sum_{j=1}^{m} \omega_j(B)\, u_{j,t-b_j} + \varepsilon_t \tag{B1.9}$$

where m inputs are each associated with a moving average operator $\omega_j(B)$ and pure time delay b_j. While each input will have its own impulse response for this model, the autoregressive parameters δ_i will be common to each input, thus constraining each impulse response to have the same decay characteristics. A more general formulation is written as

$$q_t = \sum_{j=1}^{m} q_{j,t} + \eta_t$$

$$= \sum_{j=1}^{m} \frac{\omega_j(B)}{\delta_j(B)} u_{j,t-b_j} + \eta_t , \tag{B1.10}$$

where each input is associated with the transfer function $\omega_j(B)\delta_j^{-1}(B)B^{b_j}$.

These multiple-input formulations can prove useful not only when several measured input variables are available but also where the basic linear TF model is inadequate, and an extension to the non-linear case is required.

B1.4 IDENTIFICATION OF TRANSFER FUNCTION NOISE MODELS

B1.4.1 Transfer function models

Identification of the TF(r,s,b) model involves establishing the values of r and s used to define the orders of the polynomials $\delta(B)$, $\omega(B)$, and also the pure time delay b. An estimate of the impulse response function v(B) can help infer the values of r, s and b using a relation between the impulse response function ordinates v_j, and the process model parameters δ_j, ω_j (Box and Jenkins, 1970). Equating coefficients of B in $\delta(B)v(B) = \omega(B)B^b$ gives the required relation:

$$v_j = 0 \qquad\qquad j < b$$

$$v_j = -\delta_1 v_{j-1} - \delta_2 v_{j-2} \cdot \qquad - \delta_r v_{j-r} + \omega_{j-b} \qquad b \leqslant j < b + s$$

$$v_j = -\delta_1 v_{j-1} - \delta_2 v_{j-2} - \qquad - \delta_r v_{j-r} \qquad j \geqslant b + s.$$

Note that for $j \geqslant b + s$, v_j forms an rth order difference equation $\delta(B)v_j = 0$, with r starting values v_j, $b + s - r \leqslant j \leqslant b + s - 1$. The ordinates of the impulse response function consequently provide the following information to identify b, s and r of the process model:

(i) The first b ordinates will be zero i.e. $v_j = 0$, $0 \leqslant j < b$.

(ii) The next (s-r) values follow no fixed pattern i.e. v_j for $b \leqslant j \leqslant b + s - r - 1$. These irregular ordinates will be absent if $r > s - 1$.

(iii) The remainder follow an rth order difference equation i.e. v_j for $j \geqslant b + s - r$ with r starting values v_j, $b + s - r \leqslant j \leqslant b + s - 1$

These characteristics of the impulse response of a TF(r,s,b) model can be used to help identify values of r, s and b from an estimated impulse response function.

An approximate technique to obtain an estimate of the impulse response function is based on a cross-correlation analysis between flow and rainfall; the cross-correlation function itself is of little use since autocorrelation of the separate rainfall and flow sequences in general leads to spurious cross-correlations. However if the rainfall sequence is first 'prewhitened' by identifying and estimating a stochastic rainfall model, then this model can be used to convert the rainfall sequence to a residual white noise (uncorrelated) sequence, $\alpha_t = \theta^{-1}(B)\phi(B)p_t$. This same model is then used to transform flow to a sequence $\beta_t = \theta^{-1}(B)\phi(B)q_t$, which in general will not be white noise. The cross-correlation function, $\rho_{\alpha\beta}(.)$, between the prewhitened series α_t, β_t can be shown to be proportional to the impulse response function, v(B), such that

$$v_k = \frac{\sigma_\beta}{\sigma_\alpha} \rho_{\alpha\beta}(k) \qquad k = 0, 1,$$

where σ_α, σ_β are the standard deviations of α_t and β_t.

B1.4.2 ARMA(p,q) Noise Models

The order (p,q) of the stochastic model to be used for the noise series n_t is first found by examining the sample autocorrelation and partial autocorrelation functions. Whereas the autocorrelation function indicates the correlation between variables within a time series at different time lags, the partial autocorrelation function indicates the correlation remaining after the linear dependence on variables at intervening time lags has been removed. Consequently for a pure autoregressive process the partial autocorrelation function dies out at lags beyond the order, p, of the autoregressive process; similarly the autocorrelation function of a pure moving average process dies out at lags beyond the order, q, of the process. These correlation functions are therefore particularly useful in identifying the model order of pure processes, and with experience can also be helpful in identifying mixed ARMA processes of low order.

B1.5 Parameter estimation for transfer function noise models

The Instrumental Variable - Approximate Maximum Likelihood (IVAML) Algorithm (Young et al., 1971) is applicable to systems where the output variable is comprised of the sum of a deterministic component and a stochastic component as in equation (B1.1). The stochastic component may be attributed wholly to measurement noise in q_t, or might include the disturbance effect of model errors. However, the input variables are treated as deterministic or noise free. If equation (B1.8) is used as a basis for parameter estimation, and measurement and parameter vectors are defined as

$$x_t^T = (-q_{t-1}, -q_{t-2}, \dots -q_{t-r}, p_{t-b}, p_{t-b-1}, \dots p_{t-b-s-1})$$

$$\theta_t^T = (\delta_1, \delta_2, \dots \delta_r, \omega_0, \omega_1, \dots \omega_{s-1}),$$

then equation (B1.8) can be written as

$$q_t = x_t^T \theta_t + \varepsilon_t \qquad \text{(B1.11)}$$

The presence of stochastic disturbances in the elements of x_t results in the noise ε_t being autocorrelated; as a result ε_t will also be cross correlated with q_{t-1}, q_{t-2}, ... which make up the measurement vector x_t. This cross-correlation is the origin of the inconsistent least squares estimates. If, on the other hand, equation (B1.1) is used as the basis for the estimation, then the resulting model is

$$q_t = x_t^T \theta_t + \eta_t \qquad \text{(B1.12)}$$

where

$$x_t^T = (-q_{t-1}, -q_{t-2}, \ldots -q_{t-r}, p_{t-b}, p_{t-b-1}, \ldots p_{t-b-s-1})$$

As the elements of x_t are deterministic variables uncorrelated with η_t, this model provides a basis for consistent least squares estimates of the elements of θ_t. However, the elements q_{t-1}, q_{t-2} of x_t are unknown; to overcome this problem, an estimate of x_t, denoted by $\hat{x}_t$ and referred to as an instrumental variable (IV) vector, can be provided which is defined to be highly correlated with x_t but uncorrelated with the noise η_t.

The instrumental variable vector is generated from a linear transfer function model as follows. Firstly, the estimation of the parameter vector θ is formulated in recursive form so that the estimate is updated at each time point. After all the available data have been processed, an estimate of θ is available; this can then be used to generate a set of estimates of q_t recursively as

$$\hat{q}_t = \hat{x}_t^T \hat{\theta} \qquad \text{(B1.13)}$$

This set of estimates is then used in conjunction with the (assumed) deterministic rainfall input as the IV variable for the next 'pass' through the data i.e.

$$\hat{x}_t^T = (-\hat{q}_{t-1}, -\hat{q}_{t-2}, \ldots -\hat{q}_{t-r}, p_{t-b}, p_{t-b-1}, \ldots p_{t-b-s-1}) \qquad \text{(B1.14)}$$

and this procedure is repeated until stability is achieved in the estimate of θ. Details of the recursive estimation algorithm are given in Young et al (1971) and the algorithm is summarized in Table B.1 and Figure B.2.

Once the parameter vector θ has been estimated, the parameters of a noise model of the form described in Section B1.2 are estimated using an approximate maximum likelihood (AML) method which is again formulated in a recursive form. Using the estimated parameter vector $\hat{\theta}$ and the final instrumental variable vector $\hat{x}_t$, a series of estimated residuals is generated as

$$\hat{\eta}_t = q_t - \hat{q}_t = q_t - \hat{x}_t^T \hat{\theta} \qquad \text{(B1.15)}$$

The estimation of the parameters of the noise model for η_t is approached as follows. Equation (B1.5) may be written as

TABLE B.1 Summary of Recursive Instrumental Variable Algorithm

MODEL	
System equation	$\theta_{t+1} = \theta_t$
Measurement equation	$q_t = \mathit{q}_t + \eta_t = x_t^T \theta_t + \eta_t$
ALGORITHM	
One-step ahead forecast	$q_{t\|t-1} = x_t^T \hat{\theta}_{t-1}$
Innovation (1-step ahead forecast error)	$a_{t\|t-1} = q_t - q_{t\|t-1}$
Variance of innovation error	$\sigma^2_{t\|t} = \sigma^2_\eta + x_t^T P_{t-1\|t-1} x_t^*$
Kalman gain	$K_t = P_{t-1\|t-1} x_t^* \sigma^2_{t\|t}$
Parameter estimate update	$\hat{\theta}_t = \hat{\theta}_{t-1} + K_t a_{t\|t-1}$
Variance - covariance matrix of parameter estimation error	$P_{t\|t} = (I - K_t x_t^T) P_{t-1\|t-1}$
Instrumental variable vector	$x_t^* \quad (\mathit{q}_{t-1}, \ldots, \mathit{q}_{t-r}, P_{t-b}, \ldots, P_{t-b-s-1})^T$
Explanatory variable vector	$x_t = (q_{t-1}, \ldots, q_{t-r}, P_{t-b}, \ldots, P_{t-b-s-1})$
AUXILIARY MODEL	
IV estimate	$\mathit{q}_t = x_t^{*T} \hat{\theta}$
Noise series estimate	$\hat{\eta}_t = q_t - \mathit{q}_t$

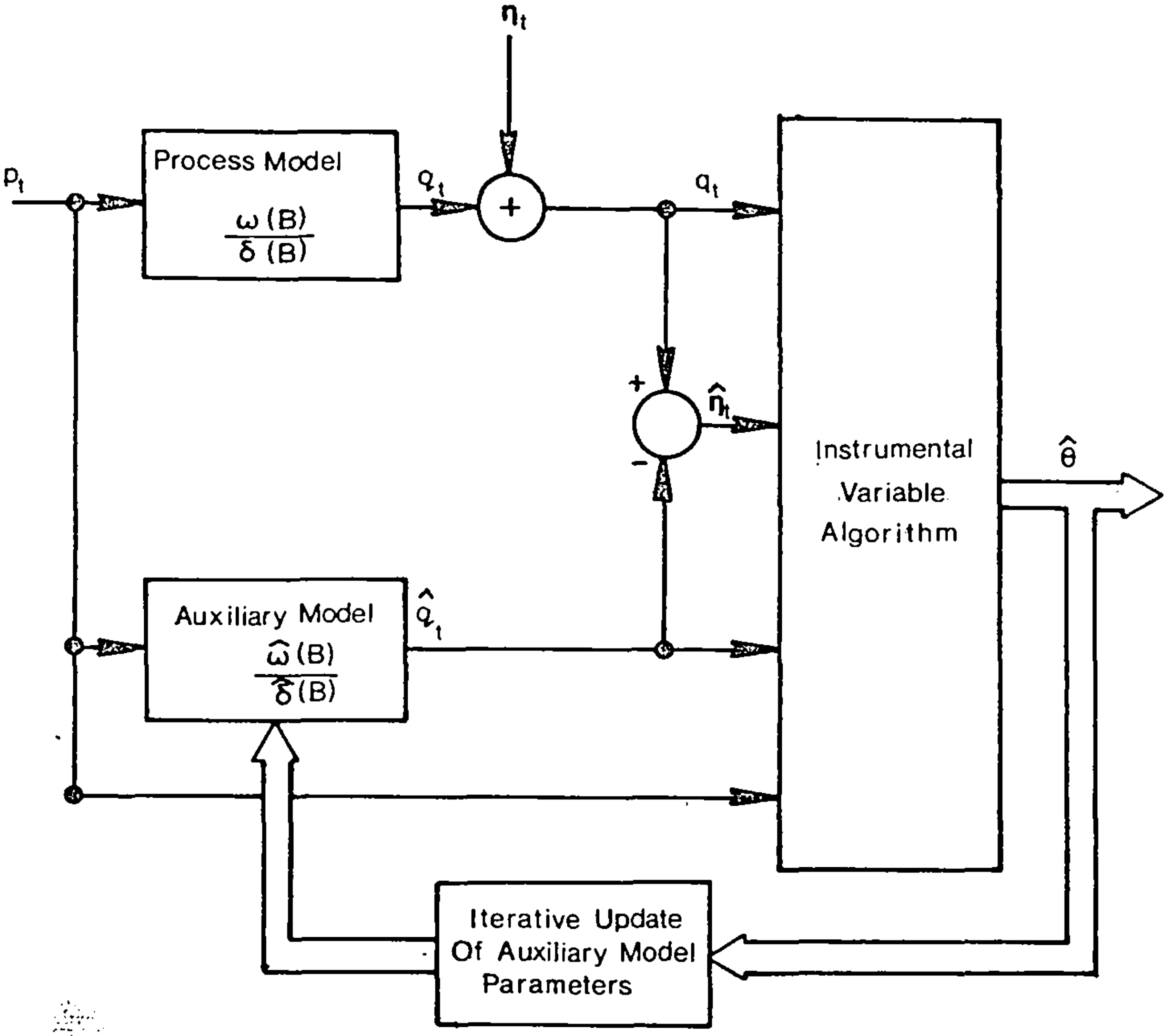


Figure B.2 The instrumental variable approach to parameter estimation

$$\eta_t = \xi_t^T \beta + a_t \qquad \text{(B1.16)}$$

where ξ_t is an explanatory variable vector defined as

$$\xi_t = (-\eta_{t-1}, -\eta_{t-2}, \ldots -\eta_{t-p}, a_{t-1}, \ldots a_{t-q})^T \qquad \text{(B1.17)}$$

and

$$\beta \quad (\phi_1, \phi_2, \ldots \phi_p, \theta_1, \theta_2, \ldots \theta_q)^T \qquad \text{(B1.18)}$$

However, the terms $\eta_{t-1}, \eta_{t-2}, \ldots, a_{t-1}, a_{t-2}, \ldots$ in equation (B1.16) are unknown; estimates of η_t are obtained from (B1.15) while estimates of a_t may be obtained using (B1.16) as

$$a_{t|t-1} = \hat{\eta}_t - \hat{\xi}_t^T \hat{\beta} , \qquad \text{(B1.19)}$$

where $a_{t|t-1}$ denotes the one-step ahead forecast error at time t using information up to time (t-1), and $\hat{\xi}_t$ is now defined as

$$\hat{\xi}_t = (-\hat{\eta}_{t-1}, -\hat{\eta}_{t-2}, \ldots -\hat{\eta}_{t-p}, \hat{a}_{t-1|t-2} \ldots a_{t-q|t-q-1})^T . \qquad \text{(B1.20)}$$

A recursive least squares algorithm can then be applied to yield *consistent estimates of the parameter vector β; initially the* explanatory variable vector is defined as (assuming $p \geqslant q$)

$$\hat{\xi}_{p+1} = (-\hat{\eta}_p, -\hat{\eta}_{p-1}, \ldots, -\hat{\eta}_1, 0, \ldots , 0) \qquad \text{(B1.21)}$$

setting $a_{p|p-1}, a_{p-1|p-2} \ldots , a_{p+1-q|p-q}$ equal to their expected values of zero; equations (B1.18) and (B1.19) are then used recursively in conjunction with the least squares algorithm as summarized in Table B.2. A number of passes through the data is necessary until stability is achieved in the estimate of β.

TABLE B2 : Summary of Recursive Approximate Maximum Likelihood Algorithm

MODEL	
System equation	$\beta_{t+1} = \beta_t$
Measurement equation	$\eta_t = \xi_t^T \beta_t + a_t$
ALGORITHM	
One-step ahead forecast	$\hat{\eta}_{t\|t-1} = \hat{\xi}_t^T \hat{\beta}_{t-1}$
Innovation (1 step ahead forecast error)	$\hat{a}_t = \hat{\eta}_t - \hat{\eta}_{t\|t-1}$
Variance of innovation error	$\sigma^2_{t\|t} = \sigma^2_a + \hat{\xi}_t^T P_{t-1\|t-1} \hat{\xi}_t$
Kalman gain	$K_t = P_{t-1\|t-1} \hat{\xi}_t \sigma^{-2}_{t\|t}$
Parameter estimate update	$\hat{\beta}_t = \hat{\beta}_{t-1} + K_t \hat{a}_t$
Variance-covariance matrix of parameter estimation error	$P_{t\|t} = (I - K_t \hat{\xi}_t^T) P_{t-1\|t-1}$
Explanatory variable vector	$\hat{\xi}_t = (\hat{\eta}_{t-1}, \ldots, \hat{\eta}_{t-p}, \hat{a}_{t-1}, \quad , \hat{a}_{t-q})^T$

B.2 CONSTRAINED LINEAR SYSTEM (CLS) MODELS

B.2 THE CONSTRAINED LINEAR SYSTEMS (CLS) MODEL

B.2.1 Linear modelling using CLS

The basis of the CLS model is a multiple-input single-output linear system which has been developed for hydrological application by Natale and Todini (1976a, b) and has been applied in non-linear form to daily rainfall-runoff modelling by Todini and Wallis (1977). In linear form, the model is written as

$$\underline{q} = \underline{U}\,\underline{v} + \underline{\varepsilon} \tag{B.2.1}$$

where $\underline{q}$ is an $(N \times 1)$ vector of discrete outputs (streamflow) sampled at a time interval Δt, $\underline{U}$ is an $(N \times nk)$ partitioned matrix of discrete time input vectors, $\underline{v}$ is an $(nk \times 1)$ vector of impulse responses, and n and N are respectively the number of inputs and the number of concurrent observations on each input and the output. Usually the estimate of $\underline{v}$ is obtained through a straightforward application of least squares involving the inversion of the matrix $(\underline{U}^T \underline{\Sigma}_{\varepsilon}^{-1}\ \underline{U})$ where $\underline{\Sigma}_{\varepsilon}$ is the variance-covariance matrix of the errors. However, this approach has a number of disadvantages, among which are

(i) the matrix $(\underline{U}^T\ \underline{\Sigma}_{\varepsilon}^{-1}\ \underline{U})$ is frequently ill-conditioned (Abadie, 1970), and errors introduced through matrix inversions may introduce errors comparable to the values of the parameters to be estimated;

(ii) the estimated impulse responses may be oscillatory with a large proportion of negative values, which is in conflict with physical principles;

(iii) continuity is not necessarily maintained.

Natale and Todini (1976a, b) have developed estimation procedures for the impulse responses which do not have shortcomings (i), (ii) and (iii). Their formulation of the problem is to minimise the functional

$$J(\underline{\varepsilon}^T\underline{\varepsilon}) = \tfrac{1}{2}\,\underline{v}^T\,\underline{U}^T\,\underline{\Sigma}_{\varepsilon}^{-1}\,\underline{U}\,\underline{v} - \underline{v}^T\,\underline{U}^T\underline{\Sigma}_{\varepsilon}^{-1}\,\underline{q} \tag{B.2.2}$$

subject to the constraints that

$$\underset{\sim}{v} \geq o \tag{B2.3}$$

$$\underset{\sim}{G}\, \underset{\sim}{v} = 1 \tag{B2.4}$$

where the matrix $\underline{G}$ is defined to maintain continuity and/or account for losses in converting rainfall into runoff. The minimization of $J(\underline{\varepsilon}^T\underline{\varepsilon})$ subject to the above constraints is achieved through quadratic programming. In the above description, the lengths of the impulse response vectors have been assumed equal for ease of presentation; no essential difficulty is encountered with non-equal values of k.

To illustrate the application of the CLS model, two examples will be considered. The first involves the case where rainfall is measured at 5 gauges within a catchment (Figure B.3); it is assumed that it is required to treat these as separate inputs, and to apply the constraints (B.2.3) and (B.2.4) to the estimation of the vector of impulse responses $\underline{v}$. In this case, $\underline{G}$ is a (1 x 5k) matrix with elements

$$g_{1,(i-1)\,k\,+\,j} = \frac{\sum_{t=1}^{N-j+1} p_t^i}{\sum_{t=1}^{N} q_t} \begin{cases} i = 1, 2, \ldots, 5 \\ j = 1, 2, \ldots k \end{cases} \tag{B2.5}$$

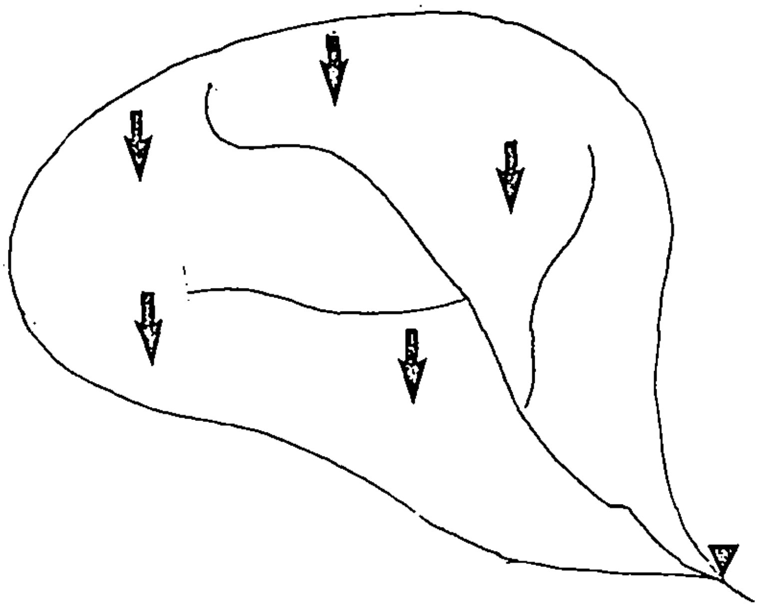
to represent the following constraint:

$$\sum_{i=1}^{n} \sum_{j=1}^{k} u_{(i-1)k+j} \sum_{t=1}^{N-j+1} p_t^i = \sum_{t=1}^{N} q_t \tag{B2.6}$$

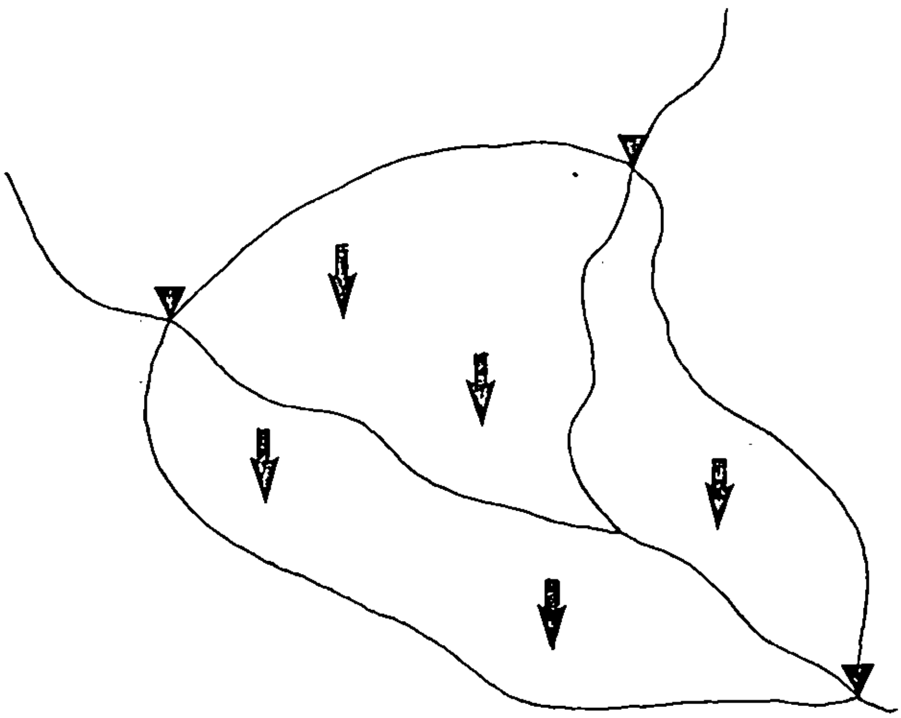
Equation (B2.6) takes into account the losses in converting the n=5 precipitation inputs p_t^i to streamflow q_t, assuming $p_t^i = 0$ for $(1 - k) \leq t \leq 0$.

The second example involves the case of m=2 upstream tributary inflows, q_t^ℓ, $\ell = 1,2$, and n-m = 7-2 = 5 precipitation inputs for the remaining contributing catchment area (Figure B3). In this case, $\underline{G}$ is an (m+1, nk) ≡ (3,7k) matrix for which the elements are all zeroes except

$$\left.\begin{aligned} g_{\ell,(\ell-1)k+j} &\equiv 1 \\ g_{3,(i-1)k+j} &= \frac{\sum_{t=1}^{N} q_t}{\sum_{t=1}^{N} [q_t - \sum_{\ell=1}^{2} q_t^\ell]} \end{aligned}\right] \begin{aligned} &\ell = 1,2 \\ &i = 3,4, \ldots 7 \\ &j = 1,2, \ldots k \end{aligned} \tag{B2.7}$$



(a) Case of n=5 precipitation inputs



(b) case of m = 2 tributary inputs and n-m = 5 precipitation inputs

Figure B3 Schematic representations of inputs to CLS model which ensures that the following continuity equation is maintained for the whole system:

$$\sum_{i=3}^{7} \sum_{j=1}^{k} v_{(i-1)k+j} \sum_{t=1}^{N-j+1} p_t^i = \sum_{t=1}^{N-j+1} \left[q_t - \sum_{\ell=1}^{2} q_t^\ell \right] \qquad \text{(B2.8)}$$

Once the parameters of the impulse response vector $\underline{v}$ have been estimated, an estimate of the a priori unknown runoff coefficient relevant to each precipitation input p_t^i is obtained as

$$\hat{\phi}_i = \sum_{j=1}^{k} \hat{v}_{(i-1)k+j} \qquad = (r+1), \ldots, n \qquad \text{(B2.9)}$$

In formulating the CLS model, it is also possible to introduce q_{t-1} as an additional input;this then results in a model of the form

$$q_t = \sum_{i=1}^{r} \delta_i \, q_{t-i} + \sum_{i=1}^{n} \sum_{j=0}^{s_i-1} \omega_j^{(i)} \, u_{t-b_i-s_i-1}^{(i)} + \varepsilon_t \qquad \text{(B2.10)}$$

which is the alternative autoregressive-moving average representation of a linear system given by Box and Jenkins (1970) with r autoregressive terms on previous outputs , s_i moving average terms and a pure time delay b_i for each input. This type of model formulation is more parsimonious (i.e. involves fewer parameters) than that given by equation (B2.1) and is particularly relevant when real-time use of the model is contemplated, as values of q_{t-1}, q_{t-2}, ... as well as $u_{t-1}^{(j)}$, $u_{t-2}^{(j)}$, ... would then be available to make forecasts of q_t, q_{t+1}, ... at time t. This then provides the model with a natural updating facility.

The CLS model can be applied to rainfall-runoff modelling, flow routing or a combination of both where the assumption of linearity is deemed reasonable. An application involving flood routing through a junction is described by Natale and Todini (1977) while Wood (1980) has used a model of the form of equation (10) for flow routing on the River Dee. The estimation of the ordinates of a unit hydrograph is an obvious application; here the use of ordinary least squares frequently results in oscillatory unit hydrographs with negative ordinates which have to be transformed into physically reasonable shape using a smoothing technique (e.g. Floods Study Report, 1975). The use of CLS obviates to a large extent the necessity for smoothing, while the constraint given by (B2.4) ensures that the unit volume criterion for the unit hydrograph is satisfied, something which is not necessarily guaranteed by smoothing.

B.2.2 Non-linear modelling using CLS

The linear form of the basic CLS model may prove restrictive for rainfall-runoff modelling applications. This was recognised by Todini and Wallis (1977) who introduced non-linearity into the model by means of a threshold mechanism applied to the rainfall input. They applied the procedure to a single lumped rainfall input, although there is no reason why the procedure could not be applied to multiple inputs. The original procedure described by Todini and Wallis (1977) has since been improved upon and is applied as follows. An antecedent precipitation index API_t is computed at time t as

$$API_t = K_t \; API_{t-1} + p_{t-1} \qquad (B2.11)$$

with

$$K_t = \bar{K} + \alpha \cos \left[\frac{2\pi}{365} (t - \phi)\right] \qquad (B2.12)$$

where $\bar{K}$, α and ϕ are parameters describing the seasonal variation in K_t. If T is then selected as a threshold value of API_t, then the following operation is performed on the input vector $u_t = p_t$ to generate two separate input vectors:

if $API_t > T$, then the value of precipitation at time t-1, p_{t-1}, is set to zero in the first input vector, and p_{t-1} is stored in the corresponding location of the second input vector;

if $API_t \leq T$, the value of p_{t-1} remains in the first input vector and a zero is placed in the corresponding location of the second input vector.

The procedure is represented schematically in Figure B.4. Thus, for one threshold, two inputs are generated from a single basic input; the multiple input capability of the basic CLS model is then utilized to derive the impulse responses for these inputs, and ultimately to derive a model output which has a non-linear relationship with the original input. The basic notion underlying the model is that different response regimes operate in a catchment in response to different states of catchment wetness, with the switch from one response to another achieved through the threshold, which introduces non-linearity into the model. Further thresholds may be applied if considered necessary although this increases considerably the number of impulse response parameters to be estimated.

The selection of values for the threshold T, and the parameters $\bar{K}$, α and ϕ describing the behaviour of K_t is done on a trial and error basis and is relatively straightforward.

The application of the CLS model with thresholds to rainfall-runoff modelling is described in Todini and Wallis (1977) and O'Connell et al (1977, 1978).

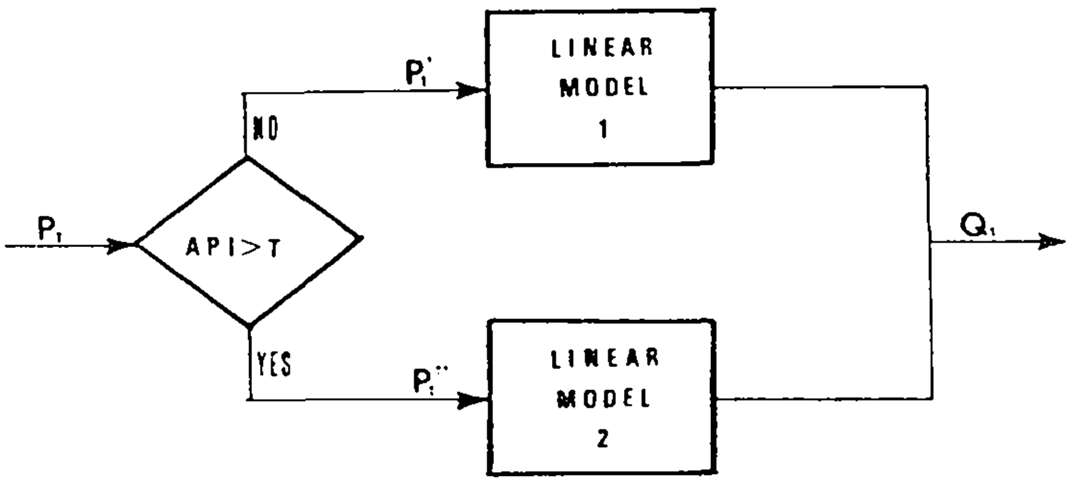


Figure B.4 CLS model with threshold

B.3 OPTIMIZATION OF PARAMETERS OF OF SIMPLE CONCEPTUAL MODELS

B.3 OPTIMIZATION OF PARAMETERS OF SIMPLE CONCEPTUAL MODELS

The algorithm used to optimize the parameters of the simple conceptual models described in Section 5 is a modified version of that developed by Rosenbrock (1960). The search geometry of the original algorithm is unchanged but modifications have been made to the way in which a minimum is found in each of the arthogonal directions.

Constraints on the variables to be optimized are introduced by applying sine-square transformations to the variables i.e. if α is a parameter which it is desired to constrain between the limits α_{max} and α_{min}, then the appropriate transformation is

$$\alpha = \alpha_{min} + (\alpha_{max} - \alpha_{min}) \, \mathrm{Sin}^2(x)$$

where x is the uncontained variable in which the search for the optimum is to be carried out using the Rosenbrock algorithm. The transformation also has the effect of reducing the parameters to be optimized to variables of the same scale.

The directions searched correspond initially to the axes of the variables. When all the directions have been searched once, new directions are defined, one of which is the direction of advance during the first iteration (i.e. the vector joining the initial and final points) and the others are orthogonal to this. New searches are made in these directions and when new minima have been estimated the directions are redefined as before and so on.

The minimum along each direction is estimated by calculating the error function at a series of points. At the start of each linear search the variable is altered by 2 per cent and the error function is computed again. If an initial failure is registered the direction of search is reversed. If a success is indicated by a decrease in the error function, the last value of the variable is altered by 3 per cent, then by 4.5 per cent, and this magnification of the steps continues until a failure is registered. The minimum is predicted from the three best error function values by quadratic interpolation using finite difference approximations; if the estimation of the minimum is found to be within a certain tolerance the next direction is searched from this point. When the function or the variables cease to change significantly a minimum is assumed to be found and the search is terminated by means of a convergence criterion.

The error function used to optimize the parameters of the simple conceptual models was the sum of squares function

$$F^2 = \Sigma(q_t - \mathit{q}_t)^2$$

where q_t denotes observed discharge and q_t denotes simulated discharge from the model.

REFERENCES

REFERENCES

ABADIE J. (1970) Integer and Non-Linear Programming, North Holland, Amsterdam, pp 544.

BOX, G. E. P., JENKINS, G. M. (1970) Time Series Analysis, Forecasting and Control, Holden Day, San Francisco, pp 553.

CRAWFORD, N., LINSLEY, R. K. (1963) A conceptual model of the hydrologic cycle, IASH Symposium on Surface Waters, Berkely.

JONCH-CLAUSEN, T. (1979) SYSTEME HYDROLOGIQUE EUROPEAN: A short description, Danish Hydraulic Institute Report, pp 17 + appendices.

MOORE, R. J., O'CONNELL, P. E. (1978) Real-time forecasting of flood events using transfer function noise models: Part 1, Report to the Water Research Centre, UK, pp 118.

MOORE, R. J. (1980) Real-time forecasting of flood events using transfer function noise models: Part 2, Report to the Water Research Centre, UK, pp 155.

NATALE, L., TODINI, E. (1976a) A stable estimator for linear models: I, Wat. Resour. Res., 12(4), 667-671

NATALE, L., TODINI, E. (1976b) A stable estimator for linear models: II, Wat. Resour. Res., 12(4), 672-676.

NATURAL ENVIRONMENT RESEARCH COUNCIL (1975) Floods Study Report, 5 Volumes, Institute of Hydrology, Wallingford, UK.

O'CONNELL, P. E. ET. AL. (1977) Methods for evaluating the UK raingauge network, IH Report No. 40, Wallingford, pp. 262.

O'CONNELL, P. E. ET. AL (1978) Rationalization of the Wessex Water Authority raingauge network, IH Report No. 51, Wallingford, pp. 179.

ROCKWOOD, D. M. (1980) Application of the COSSARR model to forecasting river and reservoir conditions in the Citarum River Basin, UNDP/WMO Project INS/78/038 Report, pp 38.

ROSENBROCK, H. H. (1960) An automatic method for finding the greatest or least value of a function, Computer Jour., 3, 175-184.

SANGSNIT, S. (1980) SSARR watershed model and its application for hydrological studies, UNDP/WMO Project INS/78/038 Report, pp 60.

SANGSNIT, S., MAUNG, O. (1981) Rainfall characteristics over the Citarum River Basin, UNDP/WMO Project INS/78/038 Report, in preparation.

SNYDER, W. N. (1955) Hydrograph analysis by the method of least squares, Proc. ASCE, Jour. Hydraul. Div., 81, 1-25.

SUGAWARA, M. (1980) Runoff analysis of the Citarum River, Indonesia, by the Tank Model, pp. 15.

TODINI, E., WALLIS, J. R. (1977) Using CLS for daily or longer period rainfall runoff modelling, in 'Mathematical Models for Surface Water Hydrology' (Eds. T.A. CIRIANI, U. MAIONE, J. R. WALLIS) Wiley, London, 149-168.

WOOD, E.F. (1980) An adaptive, input-output flow routing model, Proc.1st International Workshop on Real-time Hydrological Forecasting and Control, (Ed. P.E. O'Connell), July, 1977, Institute of Hydrology, Wallingford, pp.204.

YOUNG, P. C., SHELLSWELL, S. H., NEETHLING, C. G. (1971) A recursive approach to time series analysis, CUED/B-CONTROL/TRK, Dept. of Engineering, University of Cambridge, pp. 69.